UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

WASHINGTON, DC 20460

October 2, 2002

OFFICE OF
ENVIRONMENTAL INFORMATION

MEMORANDUM

SUBJECT: Review of *Guidance on Systematic Planning for Environmental Data Collection Using Performance and Acceptance Criteria (EPA QA/G-4A)*

FROM: Nancy W. Wentworth */s/ Nancy W. Wentworth*
Director, Quality Staff (2811R)

TO: Peer Review Panel

Attached is the Peer Review Draft *Guidance on Systematic Planning for Environmental Data Collection Using Performance and Acceptance Criteria (EPA QA/G-4A)*. Your help in supplying review comments would be greatly appreciated. This document is intended to assist a wide audience of environmental analysts, managers, and decision makers in using systematic planning for the collection of environmental data.

This document provides guidance on how to apply systematic planning for data collection using performance and acceptance criteria. The use of systematic planning is mandated through EPA Order 5360.1 A2 (May 2000) and implemented through the addressing of a set of planning elements. These elements may be addressed through application of the Data Quality Objectives (DQO) Process. When the project's objective is to choose between two opposing conditions (e.g. regulate/not regulate) the recommended approach is to apply the full DQO Process which is described in *Guidance on the Data Quality Objectives Process EPA/QA G-4*. When the project's objectives do not appear to involve decision making, recourse to related criteria must be made. This can be done through application of the Performance and Acceptance Criteria (PAC) Process, a methodology that complements the DQO Process when a clear decision point is not required.

This guidance shows how to use the PAC Process and consists of an introduction, an outline of systematic planning, the seven steps of the PAC process, some examples of PAC, case studies, and finally how this all fits into the overall collection of environmental data. As a reviewer of this document, please give attention to the organizational aspects of this guidance as well as the discussion aspects of PAC in the document. Some specifics the reviewer could consider include:

Chapter 1: The DQO Process is well established and used successfully in environmental investigations, and yet project leaders seem reluctant to tailor the DQO Process to situations where decision making is not the primary objective. Does the PAC Process make this easier? Does the chapter make clear that the PAC Process is a variant of the DQO Process? Does the chapter make clear that the PAC Process is not a replacement of the DQO Process? Does Table 1 adequately show the commonalities between the two Processes and the elements of systematic planning?

Chapter 2: The intent of this chapter is to lay the groundwork of systematic planning, is it adequate?

Chapter 3: This chapter discusses the steps of the PAC Process. In order to keep the strong connection with the DQO Process, we elected to make the PAC Process parallel the seven steps of the DQO Process, does this work? Does the list of activities and outputs agree with the intent of the step? Can you add to these activities and outputs?

Chapter 4: The intent here was to "whet the appetite" by showing examples of possible applications of the PAC Process, is this effective? Can you improve on the substance of the criteria? Is it clear that Acceptance Criteria apply to existing data for possible inclusion in the study, and Performance Criteria apply to data that will be generated for the study? How could these be improved?

Chapter 5: The case studies are simplified versions of actual studies performed in the environmental field; they were not intended to be templates for possible use. Are the case studies clear in how systematic planning was used and PAC applies? Should the PAC Process be more statistical in nature? We have used the technique of giving background information to provide information at each step of the PAC Process, is this sufficient or should the information be incorporated into the general text?

Chapter 6: This chapter is simply to remind the user of the three phases of a project and the necessity of having an approved QA Project Plan before collecting data, should this chapter be expanded?

Please feel free to offer comments and suggestions that go beyond this charge, as you see fit. Use the line numbers provided in the document to reference specific sections with recommended changes. All comments are requested from minor typographical elements to major changes in direction; they may be in hard copy or submitted electronically. I appreciate your assistance in this review and would appreciate your comments by December 6, 2002. Please send written comments to:

John Warren (2811R)
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Phone: (202) 565-8531
Fax: (202) 565-2441
E-mail: warren.john@epa.gov

Attachment

Guidance on Systematic Planning for Environmental Data Collection Using Performance and Acceptance Criteria

EPA QA/G-4A

Quality Staff
Office of Environmental Information
United States Environmental Protection Agency

Washington, DC 20460

PEER REVIEW DRAFT

October 2002

FOREWORD

The U.S. Environmental Protection Agency (EPA) has developed an Agency-wide program of quality assurance (QA) for environmental data. In particular, EPA Order 5360.1 requires that all EPA organizations follow a systematic planning process to develop acceptance or performance criteria for the collection, evaluation, or use of environmental data. This guidance document, *Guidance on Systematic Planning for Environmental Data Collection Using Performance and Acceptance Criteria*, describes one approach for conducting systematic planning.

This document provides guidance to EPA program managers and planning teams. It does not impose legally binding requirements and may not apply to a particular situation based on the circumstances. EPA retains the discretion to adopt approaches on a case-by-case basis that differ from this guidance where appropriate. EPA may periodically revise this guidance without public notice.

This document is one of the *U.S. Environmental Protection Agency Quality System Series* documents. These documents describe the EPA policies and procedures for planning, implementing, and assessing the effectiveness of the Quality System. This document is valid for a period of up to five years from the official date of publication. After five years, this document will be reissued without change, revised, or withdrawn from the *U.S. Environmental Protection Agency Quality System Series* documents. Questions regarding this document or other *Quality System Series* documents should be directed to the Quality Staff at:

U.S. EPA
Quality Staff (2811R)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Phone: (202) 564-6830
Fax: (202) 565-2441
E-mail: quality@epa.gov

Copies of the *Quality System Series* documents may be obtained from the Quality Staff directly or by downloading them from its Home Page:

www.epa.gov/quality

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

After reading this chapter you should understand the structure and function of EPA's Quality System, the kinds of programs that are a part of this System, and the benefits of using systematic planning with the Performance and Acceptance Criteria Process.

Effective data collection is rarely achieved in a haphazard fashion; some form of planning for data collection has to be used. The hallmark of all good projects, studies, and decisions is a planned data collection. When a study is being planned based on existing data and information sources or involving new data collection, and where the results of the study are not clearly linked to a regulatory decision, the Agency recommends the Performance and Acceptance Criteria (PAC) Process. When data are being used to make a decision, determine an action (e.g., compliance or non-compliance with a standard), or test a statistical hypothesis, the Agency recommends the Data Quality Objectives (DQO) Process. Both of these systematic planning processes are an integral part of the EPA's Quality System. The present document discusses systematic planning using performance and acceptance criteria. For information on the Data Quality Objectives Process, see *Guidance for the Data Quality Objectives Process*, *EPA QA/G-4* (EPA, 2000a).

Who can use this document? This guidance is intended for project managers, researchers, analysts, technical staff, stakeholders, and others wishing to use systematic planning to guide data collection efforts to ensure defensible data with measurable quality characteristics.

1.1 EPA QUALITY SYSTEM REQUIREMENTS

EPA Order 5360.1 A2 (EPA, 2000b) and the applicable federal regulations establishes a mandatory Quality System that applies to all EPA organizations and organizations funded by EPA. Organizations must ensure that data collected for the characterization of environmental processes and conditions are of the appropriate type and quality for their intended use, and that environmental technologies are designed, constructed, and operated according to defined expectations. Systematic planning is a key project-level component of the EPA Quality System (Figure 1).

EPA policy is based on the national consensus standard, ANSI/ASQC E4-1994, *Specifications and Guidelines for Environmental Data Collection and Environmental Technology Programs,* which was developed by the American National Standards Institute (ANSI) and the American Society for Quality Control (ASQC), now the American Society for Quality.

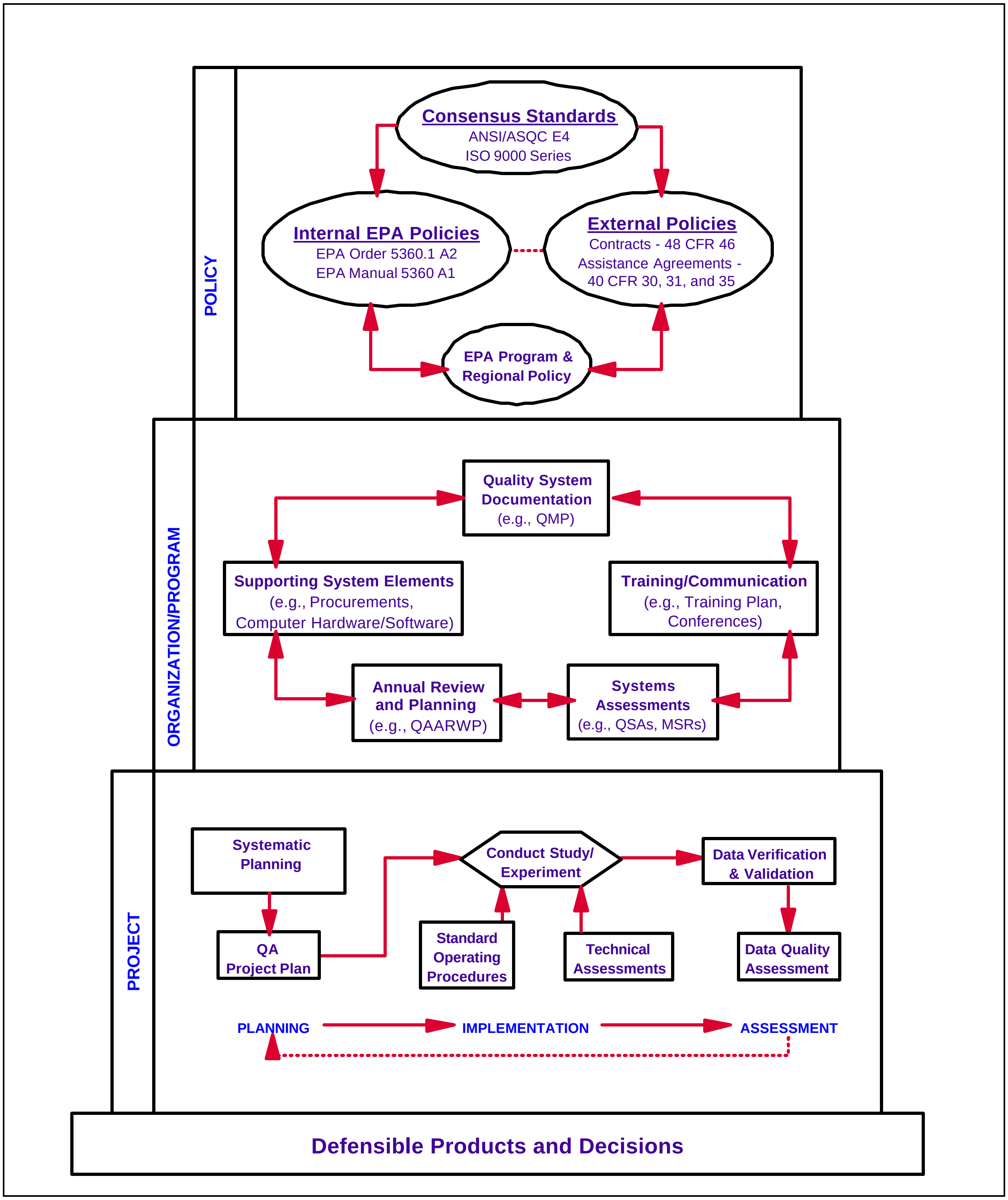

Figure 1. EPA Quality System Components

ANSI/ASQC E4-1994 describes the necessary management and technical elements for developing and implementing a quality system using a tiered approach. The standard recommends first documenting each organization-wide quality system in a Quality Management Plan or Quality Manual (to address requirements of *Part A: Management Systems* of the standard), and then documenting the applicability of the quality system to specific technical efforts in a Quality Assurance Project Plan or similar document (to address the requirements of *Part B: Collection and Evaluation of Environmental Data* of the standard). EPA has adopted this tiered approach for its mandatory Agency-wide Quality System. This document addresses Part B requirements of the standard for systematic planning for environmental data collection.

In accordance with Order 5360.1 A2, EPA requires that:

- Environmental programs performed for, or by, the Agency be supported by data of the type and quality appropriate to their expected use. Environmental data includes information collected directly from measurements, produced from models, and compiled from other sources such as databases or the literature.

- Decisions involving the design, construction, and operation of environmental technology be supported by appropriate quality-assured engineering standards and practices. Environmental technology includes treatment systems, pollution control systems and devices, waste remediation, and storage methods.

Quality specifications for non-EPA organizations are defined in the Code of Federal Regulations and information on how to satisfy these for organizations receiving financial assistance from the EPA may be found in *EPA Requirements for Quality Management Plans (QA/R-2)*, and *EPA Requirements for QA Project Plans (QA/R-5)*.

Specifications for EPA Organizations are to be found in the *EPA Quality Manual for Environmental Programs* (EPA, 2000c), which defines requirements for implementing EPA's Quality System. The Order defines the quality requirements, and the Manual presents the mandatory "how to" for implementing some of these requirements.

EPA's Quality System (Figure 1) comprises three levels – policy, organization/program, and individual project:

- *Policy:* This level addresses Agency-wide quality policies and regulations that both EPA organizations and non-EPA organizations must address.

- *Organization/Program:* This level addresses the management and implementation component of the individual Quality System.

- *Project:* This level addresses the specific components that are applied to individual projects to ensure that the needs of the organization are met.

EPA has developed a *Quality System Series* of documents that provide guidelines to help organizations ensure that data collected for the characterization of environmental processes and conditions are of the appropriate type and quality for their intended use. Documents useful in planning for data collection include:

- *EPA Requirements for Quality Management Plans (EPA QA/R-2)*,
- *Guidance for the Data Quality Objectives Process (EPA QA/G-4)*,
- *Decision Error Feasibility Trials (DEFT) Software for the Data Quality Objectives Process (EPA QA/G-4D)*,
- *Guidance for the Data Quality Objectives Process for Hazardous Waste Sites (EPA QA/G-4HW)*,
- *EPA Requirements for QA Project Plans (EPA QA/R-5)*,
- *Guidance on Quality Assurance Project Plans (EPA QA/G-5)*,
- *Guidance for Choosing a Sampling Design for Environmental Data Collection (EPA QA/G5S),*
- *Guidance on Data Quality Indicators (EPA QA/G5i),*
- *Guidance for the Preparation of Standard Operating Procedures for Quality-Related Documents (EPA QA/G-6)*,
- *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (EPA QA/G-9)*, and
- *Data Quality Evaluation Statistical Toolbox (DataQUEST) (EPA QA/G-9D)*.

1.2 SYSTEMATIC PLANNING FOR PROJECT DEVELOPMENT

EPA Order 5360.1 A2 (2000c) requires that all EPA organizations follow a systematic planning process to develop acceptance or performance criteria for the collection, evaluation, or use of environmental data. A systematic planning process is the first component in the *planning phase* of the project tier, while the actual data collection activities are in the *implementation phase* of this tier (Figure 2).

What is systematic planning? Systematic planning is a planning process based on the scientific method and includes concepts such as objectivity of approach and acceptability of results. Systematic planning is based on a common-sense, graded approach to ensure that the level of detail in planning is commensurate with the importance and intended use of the work and the available resources. The elements of a systematic planning approach to data collection include:

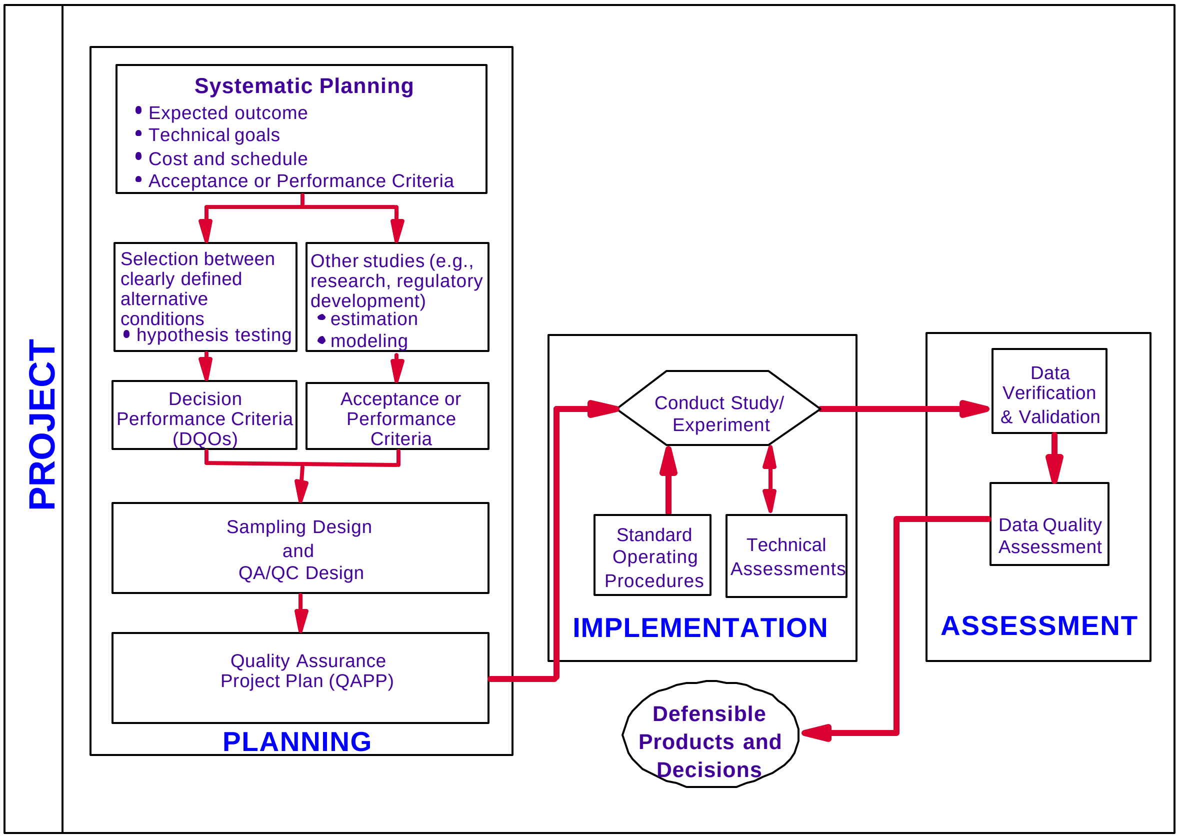

Figure 2. Systematic Planning and the EPA Quality System

- identification and involvement of the project manager, sponsoring organization and responsible official, project personnel, stakeholders, scientific experts, etc. (e.g., all customers and suppliers);

- description of the project goal, objectives, and questions and issues to be addressed;

- identification of project schedule, resources (including budget), milestones, and any applicable requirements (e.g., regulatory requirements, contractual requirements);

- identification of the type of data needed and how the data will be used to support the project's objectives;

- determination of the quantity of data needed and specification of performance criteria for measuring quality;

- description of how, when, and where the data will be obtained (including existing data) and identification of any constraints on data collection;

- specification of needed QA and quality control (QC) activities to assess the quality performance criteria (e.g, QC samples for both the field and laboratory, audits, technical assessments, performance evaluations, etc.); and

- description of how the acquired data will be analyzed (either in the field or the laboratory), evaluated (i.e., QA review, validation, verification), and assessed against its intended use and the quality performance criteria.

Systematic planning is conducted before the collection of data and the Agency recommend the DQO Process or its derivative, the PAC Process.

What is the DQO Process? When data are being used in direct support of a decision, the Agency's recommended systematic planning tool is the DQO Process. Frequently the DQO Process is used to plan a study that will utilize data to select between two opposing conditions (e.g., decision making or compliance). The DQO Process is an iterative seven-step planning approach to develop sampling designs for data collection activities that support decision making. This process uses systematic planning, and for many types of problems, recommends the use of statistical hypothesis testing to differentiate between two or more clearly defined alternatives. A summary of the seven steps of the DQO Process is presented as Figure 3.

What is the PAC Process? When data are being used for descriptive purposes, to generate estimates, or to support inferences that are not directly linked to a decision, the Agency's recommended systematic planning tool is the PAC Process. The PAC Process is iterative, similar to the DQO Process and shares some of the same steps; however, instead of focusing on specifying tolerable limits on decision errors, the PAC process provides a systematic tool for planning and designing studies where an Agency decision is not the intended outcome. A summary of the seven steps of the PAC Process is presented in Figure 4.

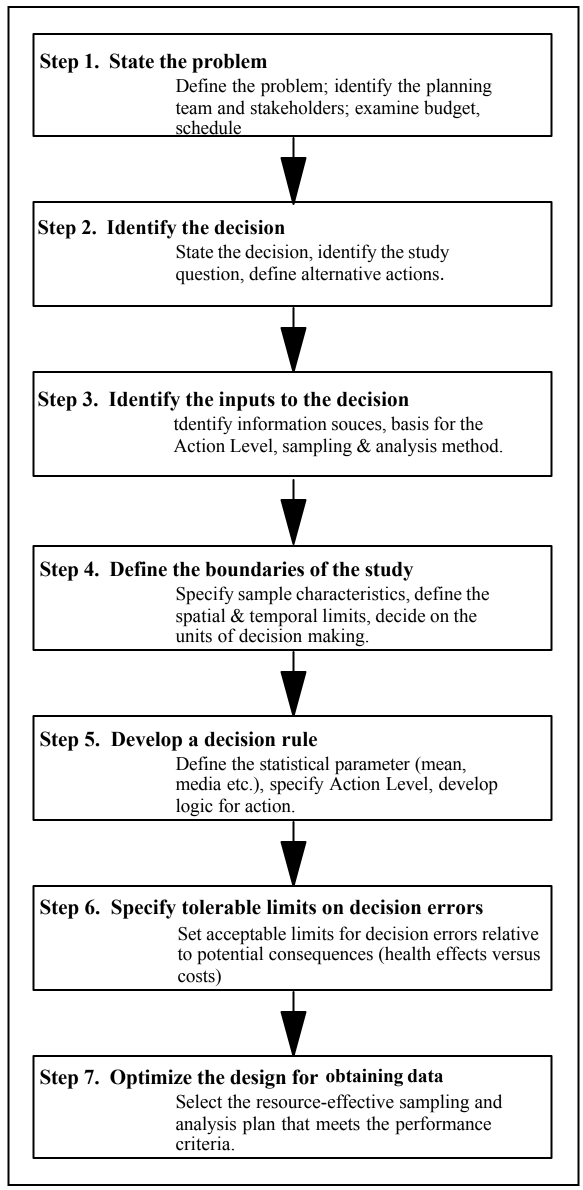

Figure 3. The Data Quality Objectives Process

Is it obvious how to choose between the two Processes? In general, the simple dichotomy:

Decision making	÷	full DQO Process
Estimation	÷	DQO adapted to PAC Process

allows for an easy separation. There are, however, instances where the differences are sufficiently close that choosing between the processes is not obvious. For example, in a research study the determination as to advancing to a subsequent stage of the investigation could be regarded as both DQO and PAC, the results of one stage of the research (performance measures) if acceptable, would lead to the commencement of the subsequent stage (the decision to continue having been made). When the choice between processes is not clear, the DQO Process should be used precisely tailored to the specific problem under investigation.

How do the DQO Process and PAC Process relate to systematic planning? The DQO Process and PAC Process share common elements but have different objectives; however, both cover the key aspects of systematic planning. Table 1 shows the connection between the processes and systematic planning [*EPA Quality Manual for Environmental Programs* (EPA, 2000c), Section 3.3.8).

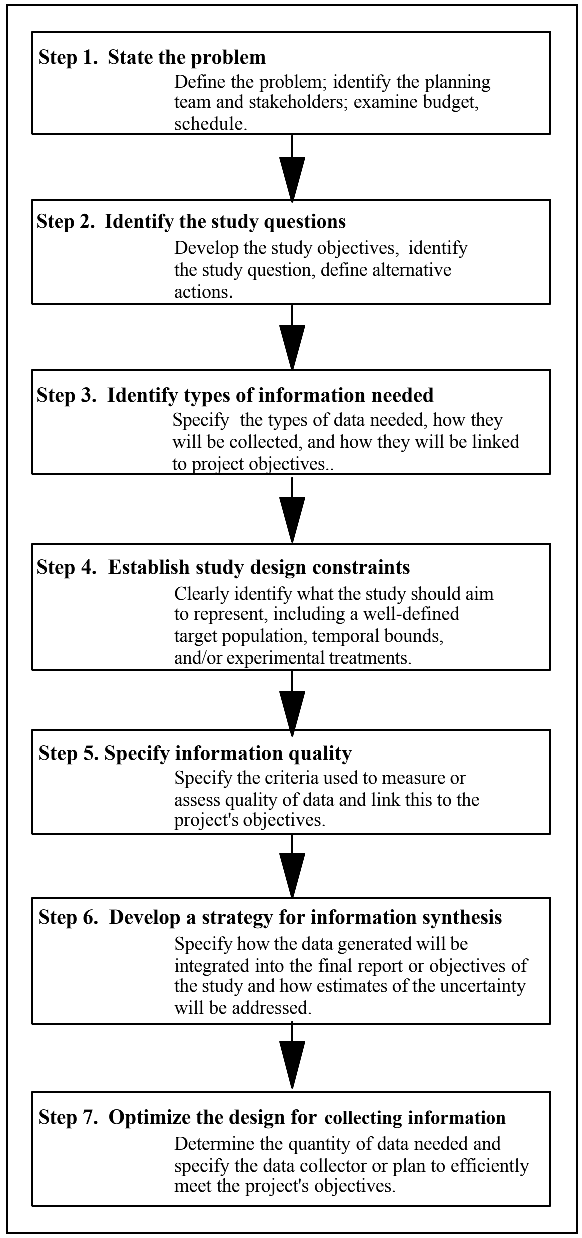

Figure 4. The Performance and Acceptance Criteria Process

Is the PAC Process a replacement of the DQO Process? No. The PAC Process may be regarded as an adaptation of the DQO Process to situations where decision making is not the primary focus of the data collection. In those cases, the PAC Process helps focus attention on the key elements of systematic planning.

Does this guidance include both the DQO Process and the PAC Process? No. The DQO Process is covered in depth in *Guidance for the Data Quality Objectives Process (EPA QA/G-4)*. The present document discusses systematic planning using the PAC Process.

Table 1. Systematic Planning and the DQO and PAC Processes

Elements of Systematic Planning	Corresponding Step in the DQO Process	Corresponding Step in the PAC Process
Identification of the project manager, sponsoring organization, staff, stakeholders, and experts	Step 1. Define the problem	Step 1. Define the problem
Description of project objectives and issues	Step 1. Define the problem Step 2. Identify the decision	Step 1. Define the problem Step 2. Identify the study questions
Identification of the project schedule, resources, milestones, and requirements	Step 1. Define the problem Step 2. Identify the decision	Step 1. Define the problem Step 2. Identify the study questions
Identification of the type of data needed and the link to the project's objectives	Step 3. Identify inputs to the decision	Step 3. Identify types of information needed Step 4. Establish study design constraints
Determination of the quantity of data needed and how this is linked to the project's objectives	Step 4. Define the boundaries of the study Step 5. Develop a decision rule	Step 5. Specify information quality Step 6. Develop a strategy for information synthesis
Description of how, when, and where the data will be obtained, together with an identification of any constraints	Step 6. Specify limits on decision errors Step 7. Optimize the design for obtaining data	Step 6. Develop a strategy for information synthesis Step 7. Optimize the design for collecting information
Specification of QA and QC activities to assess the quality performance criteria	Contained in the QA Project Plan derived from Step 7	Contained in the QA Project Plan derived from Steps 6 and 7
Description of the methods for data analysis, evaluation, and assessment against the intended use of the data and the quality performance criteria	Contained in the QA Project Plan and by the use of Data Quality Assessment	Contained in the QA Project Plan and by the use of Data Quality Assessment

Will you always develop statistical/probabilistic sampling designs for data collection if you use the PAC Process? No. Although statistically-based designs are strongly encouraged, there are some instances in which it is not realistic (e.g., a monitoring network has already been established and cannot be altered). In research data collection, it is common to consider variations on design of experiments, a statistical technique beyond the scope of this document.

How should you use this guidance? You should use it as an aid in structuring planning for environmental data collection and use. It will help to organize the agenda for workgroup meetings; focus attention on key issues; and facilitate communication among technical experts, managers, stakeholders, decision makers, and the environmental community.

When should the PAC Process be used? The PAC Process should be used during the planning stage of an investigation that either requires new data collection or requires assembling and using existing data or information sources *before* the data are collected.

Is the PAC Process applicable only to large studies? No. The PAC Process applies to any study, regardless of its size. However, the depth and detail of PAC development depend on the project's complexity and objectives.

What are the outputs of the PAC Process? The PAC Process leads to the development of acceptance or performance criteria. Acceptance or performance criteria are based on the ultimate use of the data to be collected, and they define the quality of data required to meet the final objectives of the project or study.

What is a data collection design? A data collection design specifies the series of activities required to satisfy the performance and acceptance criteria. The design may include plans for sampling or direct observation, field or laboratory analyses, and QA or QC procedures or other tools for assessing the quality of the resulting data set. These QA and QC procedures are documented in the QA Project Plan.

Can existing data be used in the PAC Process? Yes, existing data can be very useful. At a minimum, it can be an important input to project formulation, helping planners to clarify study objectives, frame hypotheses, or devise a conceptual model for the project. In addition, existing data (either by itself or in combination with newly collected data) may be adequate for achieving the objectives of the study. Combining existing and new data can be a complex operation, requiring an analysis of the data sets' comparability and representativeness [see discussions of comparability and representativeness in *Guidance on Data Quality Indicators (EPA QA-G5i)* (EPA, 2002a)].

1.3 BENEFITS OF USING THE PAC PROCESS

Systematic planning using the PAC Process involves multi-disciplinary team members, and helps to channel their diverse perspectives toward a common focus on a successful project conclusion. This interaction results in a clear understanding of the project and the options available for conducting it. Organizations that have used the PAC Process have found the following:

- The structure of the PAC Process provides a convenient way to *document all the activities and outcomes*, communicate the basis for the data collection design to others, and *facilitates rapid review and approval* of the QA Project Plan.

- The PAC Process enables data users and relevant technical experts to participate collectively in data collection planning, and to specify their particular needs prior to data collection. The PAC process fosters *communication among all participants*, one of the central tenets of quality management practices.

- The PAC Process helps to focus studies by encouraging data users to *clarify vague objectives* and explicitly frame their study questions.

- The PAC Process is a planning tool that can *save resources* by making data collection operations more resource-effective. Good planning will streamline the study process and increase the likelihood of efficiently collecting appropriate and useful data.

Upon implementing the PAC Process, your environmental programs should therefore be strengthened by the process achieving:

- focused data requirements and optimized design for data collection;
- use of clearly developed work plans for collecting data in the field;
- a well documented basis for data collection, evaluation, and use;
- clearer statistical analysis of the final data; and
- sound, comprehensive QA Project Plans.

1.4 ORGANIZATION OF THIS DOCUMENT

This document contains an introductory chapter, followed by four chapters that describe systematic planning using performance and acceptance criteria, and ends with Chapter 6 that shows how outputs of the planning process are used to develop a QA Project Plan. This document is designed as a companion to *Guidance for the Data Quality Objectives Process (EPA QA/G-4)* (EPA, 2000a), and shares some of the same elements.

CHAPTER 2

SYSTEMATIC PLANNING FOR ENVIRONMENTAL DATA COLLECTION

After reading this chapter you should understand the overall structure of systematic planning using the Performance and Acceptance Criteria Process.

Policy and Program Requirements for the Mandatory Agency-Wide Quality System (EPA, 2000b) requires all EPA organizations and others involved in extramural agreements (e.g., contracts, grants, and cooperative or interagency agreements) with the Agency to use a systematic planning approach leading to development of acceptance or performance criteria for the collection, evaluation, or use of environmental data. The level of detail in systematic planning is to be commensurate with the importance and intended use of the work and the available resources. Systematic planning identifies the expected outcome of the project, technical goals, cost and schedule, and quality criteria for intermediate products such as data sets as well as for the final overall product of the project. The outputs from the seven steps of the PAC Process combine to give the optimal method for obtaining data to meet the project's objectives.

2.1 SEVEN-STEP SYSTEMATIC PLANNING PROCESS USING PERFORMANCE AND ACCEPTANCE CRITERIA

Like the DQO guidance, this guidance recommends a seven-step process of systematic planning for data collection. In accordance with *EPA Quality Manual for Environmental Programs* (EPA, 2000c), the planning is "based on a common sense, graded approach," which matches the scale of planning with the importance of the project and the intended use of the data. The seven steps recommended in this guidance can be summarized as follows:

1. *State the problem:* Describe project background, goals, and objectives; identify the project manager or principal investigator, sponsoring organization and responsible official, project personnel, stakeholders, scientific experts, etc.; and list key project constraints, for example, the project schedule and budget.

2. *Identify the study questions:* Specify the questions and issues to be addressed by the project.

3. *Establish study design constraints:* Summarize the considerations that drive requirements for data. If data are to be used to represent some spatial area or time frame, specify the target population (and the units of which it is composed) that the study will aim to represent. If data are to be used to conduct an experiment, specify the

conditions, treatments, and control(s) that will drive the experimental design. Describe any technical or practical constraints on data collection and use.

4. *Identify types of information needed:* Determine what kinds of data are to be gathered and establish that current technology will be adequate to generate the required measurements (e.g., confirm that adequately sensitive procedures are available). Confirm that all identified types of data are appropriate to support the project's objectives.

5. *Specify information quality:* Establish quantitative or qualitative statements regarding the level of certainty desired in the outcome of the study. For descriptive problems, generally qualitative statements are adequate. For estimation problems, specify the desired width of confidence intervals. Specify and justify the criteria to be used as measures of the quality of individual data or pieces of information that will be integrated into the project's final product.

6. *Develop a strategy for information synthesis:* Specify how the individual data points and other inputs will be reduced, analyzed, and combined into the final report or other products that achieve the goals and objectives of the study. Then specify the strategy that will be used to estimate the uncertainty or variability of the study conclusions.

7. *Optimize the design for obtaining information:* Specify a data collection protocol that provides sufficient quantity of data of appropriate quality at the least cost (i.e., the most resource-effective study design).

The first four of these planning steps can be considered preliminary aspects of scoping and defining the study, while the last three steps establish the "acceptance or performance criteria" that will define the quality of the study data.

2.2 WHAT ARE PERFORMANCE AND ACCEPTANCE CRITERIA?

Policy and Program Requirements for the Mandatory Agency-Wide Quality System EPA, (2000b) requires that systematic planning be used to develop "acceptance or performance criteria" for all work covered by the order. *EPA Quality Manual for Environmental Programs* (EPA, 2000c) further details the elements of a systematic planning approach and required documentation and emphasizes the "specification of performance criteria for measuring quality" in the context of QC and QA planning activities. *Guidance for the Data Quality Objectives Process (EPA QA/G-4)* (EPA, 2000a) also discusses performance and acceptance criteria, but in a decision-making context.

In general, performance and acceptance criteria are measures of data quality established for specific data quality indicators (DQIs) and are used to assess the sufficiency of collected information. This guidance document makes the following distinction between the two terms:

- *Performance criteria* address the adequacy of information that is to be collected for the project. These criteria often apply to new data collected for a specific use (also known as "primary" data).

- *Acceptance criteria* address the adequacy of existing information proposed for inclusion in the project. These criteria often apply to data drawn from existing sources (also known as "secondary" data).

For example, investigators collecting data on the concentration of lead in paint at a given building might require that the 95% confidence interval be no larger than ±20% of the mean. If unforeseen technical problems result in less precise or fewer measurements being made, the data might be rejected as not having met the *performance criterion* for new data collection. By contrast, investigators comparing two competing environmental measurement technologies might conduct an analysis of existing data sources (e.g., from literature or publications) and establish as an acceptance criterion that the data must be adequate to provide a 90% probability of finding a 25% difference (e.g., in the analytical precision of the two methods) as statistically significant. Existing data not providing the desired probability of finding a difference of this magnitude as significant might be rejected as not having met the *acceptance criteria* for secondary data use.

2.3 DEVELOPING PERFORMANCE AND ACCEPTANCE CRITERIA

The last three steps in the PAC Process establish the performance and acceptance criteria for data collection.

Specifying *information quality* (Step 5) involves specifying the planning team's targets for the quality of the results (descriptions, estimates, or other conclusions) that can be used either to develop a plan for data collection or evaluate the adequacy of existing data to support the intended use. In addition, this step may involve identifying specific DQIs that are relevant to the project and the appropriate QA and QC procedures that should be used. For each applicable DQI, quantitative or qualitative Measurement Quality Objectives (MQOs) should be established together with the QA procedures that will be used to ensure these MQOs have been achieved. In many cases, quantitative MQOs can be established only in context with the overall design, since they are linked directly to quantitative assumptions in the statistical design that is developed as part of Step 7. Therefore, iteration between Steps 5 and 7 is expected and encouraged.

Information synthesis (Step 6) is the process of combining the separate pieces of environmental data to form a coherent analysis that directly supports the project goals and objectives (e.g., meta-analysis of existing data sets, contaminant transport and human exposure modeling, statistical modeling of measured data, etc.). To help ensure sufficiency and cost-effectiveness in the study design, study planners should directly tie the data synthesis criteria to the study objectives.

Optimizing the design for obtaining information (Step 7) involves specifying the type, quality, and quantity of new or existing data that will meet project objectives in the most cost-effective manner. This step incorporates the outcomes of each previous step, and provides the basis for development of the QA Project Plan or similar document.

2.4 APPLYING THE GRADED APPROACH TO SYSTEMATIC PLANNING USING PERFORMANCE AND ACCEPTANCE CRITERIA

The "graded approach" is defined in *EPA Quality Manual for Environmental Programs* (EPA, 2000c) as "the process of basing the level of application of managerial controls applied to an item or work according to the intended use of the results and the degree of confidence needed in the quality of the results."

Based on application of the graded approach, the level of documentation and rigor associated with the systematic planning process may vary from one environmental data operation to the next. The level of planning detail and documentation may:

- correspond directly to the importance of the project to its stakeholders (e.g., the potential for environmental health risks to sensitive human and ecological populations);

- reflect the overall scope and budget of the effort (larger scope and budget generally correspond with a more extensive project, which in turn requires greater documentation); and

- be driven by the inherent technical complexity or the political profile of the project (more complex or politically sensitive projects require documentation of more technical details or nuances).

The output of all seven steps should be documented during systematic planning, but some elements of the plan may be less thoroughly documented and less rigorously defended in the planning documentation for relatively less involved or less demanding projects.

For example, a pilot study intended to estimate variability as an input to further planning might represent the less rigorous end of the spectrum of information collection activities. Study planners should document the objectives of the pilot study, what measurements are needed, and how they will be used (e.g., to determine the range and variability of concentrations within the study area). In cases such as this, existing data are rarely available to determine a statistical basis for sample sizes, therefore specifying quantitative performance criteria in Step 5 may not be of great value (a qualitative expression related to representativeness and perhaps measurement precision would suffice), but explaining how the data will be analyzed for use in planning the full study (Step 6) would be very useful. Other cases that might be associated with the low end of the graded approach include projects where planners agree that the fixed project schedule and budget are the main constraints on the required quantity and quality of data. Documentation of the seven-step planning process could be relatively brief, perhaps taking the form of a technical work plan. Using the graded approach does not mean avoiding or skipping any steps in the planning process, but it may simplify and streamline discussions related to specific steps of the process.

At the higher end of the graded approach are studies that collect and use direct chemical measurements and other types of primary data (e.g., that call for accurate and precise quantitative data) to support or inform major environmental or public health programs. For example, study planners may wish to conduct an extensive study to evaluate the efficacy of two or more promising new technologies with major potential for environmental improvement and cost savings. Such data collection activities may require rigorous consideration of performance criteria for data, along with relatively extensive documentation of the seven-step study planning process. Statistical evaluations of historical data or implementation of a pilot study to generate data upon which to base a statistical design may be pursued to support determinations of quantity and quality of data to be collected. Careful tradeoffs between the number of samples and quality of individual measurements might be considered to ensure that an optimal solution to the design is developed.

In these more rigorous settings, the planning steps involving information synthesis and the formation of study inferences would typically take into account statistical confidence and may include specifying the magnitude of a difference that is meaningful and other statistical concepts. The professional qualifications of the project staff (i.e., training, skills, and experience) or the robustness and reliability of the quantitative models employed might also be appropriate to consider and specify in the form of performance criteria in more intricate or important projects.

The remainder of this guidance document explains the seven planning steps of the PAC Process and provides several examples and case studies to further illustrate the concepts.

CHAPTER 3

THE SEVEN STEPS OF THE PAC PROCESS

After reading this chapter you should understand in greater detail systematic planning and the activities associated with each step of the Performance and Acceptance Criteria Process.

In the following section, each step of systematic planning using the PAC Process is described in greater detail. Beyond the further background discussion of each step, the specific activities to be performed and outputs to be generated are also described.

3.1 STEP 1: STATE THE PROBLEM

When a complex study is being planned, it is crucial that a well-organized, knowledgeable planning team be formed. In research studies led by a single investigator, it is important that access to literature and expertise beyond that immediately demanded be established early in the investigation. A carefully conducted systematic planning process will involve input from a planning team composed of various participants, who in some cases may represent diverse interests or concerns regarding the planned study. For example, a study might be funded by an EPA program office that has a particular interest in the study outcome and designed and performed by an EPA laboratory that has some shared and some different interests in the project. In such cases, it is important to bring both parties together to ensure that the study will provide the information needed for both perspectives. When appropriate, the planning team should be composed of cross-disciplinary experts familiar with the different technical aspects of the problem and different aspects of the technical approach for conducting the study. Recruitment of appropriate team members at this step of the process helps ensure that important details of the study are not overlooked or ignored, and that technical challenges will be addressed appropriately.

Once the planning team is established, the next and most important step is to define and concisely state the problem at hand. As the planning process evolves, the problem statement becomes the focus of the planning team members.

The available resources and deadlines should be ascertained at the outset of the project to prevent subsequent abandonment due to under/over allocation of resources at critical points in the project development. A clear statement of the planning process resources, constraints, and deadlines helps resolve conflicts by specifying the practical bounds of the best possible problem resolution.

Activities

- Identify or organize the planning team members, principal investigator, research study coordinator, or project leader who will have primary responsibility for resolving conflicts, balancing objectives, and maintaining progress. Ensure that the team includes representatives of the critical users of the information or data that will result from the study in addition to the technical experts involved in planning.

- Concisely describe the problem at hand, providing a problem statement that summarizes why the study is being conducted. In this problem statement, specify the needs of the primary and potential secondary users of the information or data that will result from the study. Develop a conceptual description of the problem and the study approach being considered with enough detail so that the appropriate data inputs can be identified.

Outputs

- Documentation of the roles, responsibilities, and contact information for the planning team members

- A concise statement presenting a conceptual description of the environmental problem and identifying the need or purpose for the study

- Specification of available resources, constraints and deadlines, including a working list of goals and milestones to provide project direction

3.2 STEP 2: IDENTIFY THE STUDY QUESTIONS

The purpose of this step is to ensure that all interested parties understand and agree on the study objectives, and in particular, the way in which the study questions are framed. Objectives reflect a general statement of the intent of a study and how that study is linked to addressing the environmental problem (or contributing to the field of science). Moving from the statement of objectives to specific and appropriate study questions is one of the most important steps in the PAC Process. The study questions should be framed so that they can be addressed by the data or information that will result from the study. The way in which the study questions are framed will be directly related to where in the hierarchy introduced in Chapter 2 the study best fits: a descriptive study, study to support estimation, or study designed to support specific inferences.

For descriptive studies, the study question will basically state what the data will be used to describe. For example, the question may simply be asking what the state of nature is in a particular location: e.g., what species of invertebrates, emergent plants and algae are present in specified locations along a watershed?

For estimation studies, the study question should include a statement of what is to be estimated based on the result. For example, the question may ask about some statistical parameter that may be of interest in addressing the study objectives. Framing the study question leads the planning team to agree upon the parameters of greatest interest. Therefore, it would not be sufficient to simply investigate what organic and inorganic air toxicants are present downwind from a smelter, but to frame the question in terms of the summary statistic to be estimated. For example, ask "What is the daily maximum concentration of hourly measurements of all detected air toxicants downwind from a smelter?" Estimation problems lend themselves to a more quantitative planning process, with the PAC Process generating a statement about the quality of the estimate desired. Estimates are frequently used to draw inferences, and the distinction between estimation and use for making inferences about a parameter can become blurred.

Framing a study question, when the results of the study are intended to support a specific preconceived theory, is an integral part of the scientific method. It is sometimes appropriate and useful to state the study questions and then translate these into specific, testable terms. For example, rather than just estimating the mean concentration of air toxicants, one may want to compare that concentration over time, between locations, before or after some new pollution control device is installed, etc. It may be sufficient to simply state, in common terms, what the planning team is interested in comparing or determining through the study. If it is clear that the data will be used in decision making, for example, determining possible non-compliance with a standard, the investigation will be better conducted using the DQO Process rather than the PAC Process.

Activities

- Develop and state your study objectives and frame the study questions in a manner appropriate to the type of study being conducted. If the study is intended to support specific preconceived theory, and the study question is framed as a statistical hypothesis, this step will require the planning team to carefully consider how to translate the study question into a hypothesis test and consider using the DQO Process. To achieve your study objectives, the planning team will collect new data (or derive data from existing sources) adequate to resolve the study questions. As some studies may also involve secondary objectives, it may be appropriate for the study team to prioritize among the objectives and decide which study questions should be considered first in developing the plan.

Outputs

- A clear statement of the study objectives and associated study questions, framed so that data can effectively be used to address the questions.

3.3 STEP 3: IDENTIFY TYPES OF INFORMATION NEEDED

The purpose of this step is to focus the study on the specific kinds of data or information to be collected from the population of interest. This exercise starts by developing a list of those data that will support the study questions defined in Step 2. This list typically includes measurements of variables of interest (e.g., chemical concentrations in environmental media, biological effects) that are directly associated with the environmental or experimental issue being addressed by your study as well as identifiers for each sampling unit (e.g., geographic location coordinates and sampling time for environmental monitoring data, or street address, name, social security number, and sampling time for questionnaire data from a person) or experimental treatment. In addition, it may include other types of ancillary information that can be economically collected to enhance the data interpretation and clarify the study conclusions. These latter types of information might or might not be collected based on schedule and budget constraints.

Activities

- List the types of information needed to address the study questions. This may include regulatory drivers, relevant benchmarks such as ecological screening values or historical estimates of the minimum significant difference, or ancillary data outside the main focus of the study.

- Specify the sources of information and the general methods for obtaining the needed information. This may include literature sources; existing databases; or new data collection activities, either by experimentation or through observational studies.

Outputs

- A specific list of information needed.

- Justification for the types of information needed.

- Confirmation that measurement technology is available to generate results with adequate sensitivity in the media of interest.

- A list of information sources.

3.4 STEP 4: ESTABLISH STUDY DESIGN CONSTRAINTS

The fourth step in the PAC process involves defining the constraints associated with the study. The major constraint to consider is what set of conditions, spatial area and/or time frame should data from the study (or existing data, or a combination of both) represent. *Guidance for the Data Quality Objectives Process (EPA QA/G-4)* (EPA, 2002a) provides an in-depth discussion of representativeness of environmental studies that can assist the planning team when fully considering this issue. If the study involves sampling a portion of the environment, it is very important to clearly define the target population for the study and to define the units of this population for which measurements or observations will be obtained. Establishing a clear understanding of the study population, as well as any practical or logistical issues that might affect the ability to sample all units of the population, facilitates the process of designing a representative study. The more clearly the target population is defined, the more likely the study design will support inferences drawn from the sample data. Data collected without a clear understanding of just what that population represents can lead to serious bias.

The target population is the entire collection of sampling units that you are observing, describing attributes of, generating quantitative estimates and/or making specific inferences from, or drawing scientific conclusions about. For example, a sampling unit might be one person from a population of people, one surface soil sample from all possible surface soil samples covering a potentially contaminated industrial site, or one unit/volume of air or water for which one or more measurements will be made. For experimental studies, the equivalent set of specifications would be a description of each treatment and the control for the experiment and the set of conditions that define them.

Activities

- Define the sampling unit with respect to the portion of the physical environment from which one or more samples may be taken: for example, the volume of soil contained in a grab sample; an individual person, animal, or plant; the volume of air obtained in a collection canister or that has moved through a filter; a single unit produced by a manufacturing process; or some specified volume of environmental media about which a measurement has meaning for the intended use of the data.

- Describe the total collection of sampling units in terms of space and time as well as key attributes that define the sampling units of interest. For studies of people, this may consist of population characteristics such as region, race, ethnicity, age, etc. The sampling units for the study may be a subset of the population for which data are required, such as all male children of African-American descent under the age of 12 living in a specific county. For studies of some environmental media, this activity involves generating a list of all sampling units that fall within the target population (e.g., set of all soil grab samples at a potentially contaminated waste site, set of all possible responses from a new analytical device, set of all possible simulation model outputs). The sampling units may be defined as the volume of media obtained from the sampling device, or they may be defined as some larger area or volume from which multiple samples can be taken. For experimental studies, describe the sets of conditions that define the treatments and control.

- Describe the population of interest for any additional data (e.g., ancillary data or meta-data) to be collected in addition to the main body of study data. Include a general description of the sources and methods for collecting ancillary data and how these data will be inspected to ensure compatibility with the proposed major data collection. This information will be used in Step 7, when a final information collection protocol will be designed.

Outputs

- A clear definition of the population of interest.

- Important characteristics of this population that should be accounted for in the sampling design.

- A clear definition of the sampling units that compose the population of interest, including characteristics and spatial and temporal boundaries that identify or define these units.

3.5 STEP 5: SPECIFY INFORMATION QUALITY

The PAC process leads to a set of specifications regarding the type and quality of data needed to support the intended use. These quality specifications are either (a) "performance criteria," in the case of original or primary data or (b) "acceptance criteria," in the case of existing or secondary data. Performance criteria specifications together with the appropriate level of QA practices are used during the planning phase to guide the design of the data collection effort. Acceptance criteria are used to guide the plan for evaluating existing data. After data are generated (or assembled from existing sources), both performance and acceptance criteria can be used to assess data adequacy, i.e., to determine if the description, estimate, or concept can be performed or tested with a desired level of confidence.

For measured, quantitative information, performance criteria can be used as a basis for establishing MQOs for specific data quality indicators. DQIs are quantitative and qualitative measures of principal quality attributes that include precision, bias, representativeness, comparability, completeness, and sensitivity. Typically, quantitative MQOs for the sampling and measurement components of total variability and for sensitivity are established as part of a statistical design. In addition, qualitative statements regarding representativeness are developed as part of Step 4 of the PAC Process, and a discussion of data comparability often accompanies acceptance criteria, especially when study results will be based on more than one source of data.

When a study calls for use of secondary (existing) sources of data, acceptance criteria apply. By developing and using acceptance criteria, a layer of objectivity is placed on the process—by first considering what population existing data should adequately represent to support the new study objectives, and then determining how much uncertainty is acceptable. The total uncertainty can often be further broken down into components that can be evaluated during the study. For example, it should be possible to determine, for measurement data, if the analytical detection limits were adequately low, and by looking at a number of different QC samples, whether measurement variability and bias were under control and adequate to serve the new planned use of the data.

DQIs specify the ways in which data quality will be assessed for a particular test or measure, and MQOs specify the thresholds that define minimally acceptable data. Even for less quantitative or purely qualitative information (e.g., descriptive studies), the concept of DQIs and MQOs still applies; however, the degree to which the MQOs are quantitatively linked to the performance or acceptance criteria diminishes.

Activities

- Evaluate the potential consequences of uncertainty from your study. The evaluation may differ between descriptive studies and those of a more quantitative nature such as estimation. Consider how the outcome of the study will be used, and how an under- or over-estimate of some variable of interest might affect the outcome of the study. It is often important to seek input from the full range of potential data users to assist in establishing limits on uncertainty, since different uses of the data may be affected differently. One effective way to elicit this information is to write down specific scenarios and consider the consequences should the estimation process lead to the wrong interpretation (due to unexpected high variability in data obscuring results, or because insufficient data were collected, for example). The consequences should be converted to a common metric (costs, for example) and compared so that the information collection design can be refined if possible.

- For each study question, indicate the level of uncertainty permissible. The method of presenting the uncertainty will depend on the form of the problem. The actual value selected is derived from a desire to avoid the consequences discussed above and is linked to some statement of the magnitude of the error. For an estimation problem, these quality statements will typically be characterized by an uncertainty interval width associated with a confidence level or probability. A descriptive problem will include qualitative statements about the desire to adequately represent the population of interest.

- Through the use of MQOs, specify the criteria for the measure of quality you have chosen. For a confidence interval, give an acceptable numerical value for the width of the interval and an acceptable probability of including the true value in that interval. For a tolerance interval, give an acceptable numerical value for the width of the interval and an acceptable probability of covering the desired percentile of the target population.

- Especially in studies for which the only feasible criteria are qualitative, relate the criteria to the goals of the project.

Outputs

- Specification of performance and acceptance criteria as appropriate to ensure the adequacy of the design for sample collection and measurement

3.6 STEP 6: DEVELOP A STRATEGY FOR INFORMATION SYNTHESIS

In this planning step, the planning team focuses on specifying how the various data inputs will be reduced and used to address the study questions framed in Step 2. Given the range of project types for which performance and assessment criteria apply, this step will vary from a simple discussion of how data will be summarized and used to describe some state of nature, to a complex discussion of how data will be combined into a framework, structure, model, or set of models for achieving the goals and objectives of the study. Between these extremes, this step might specify how data will be evaluated and used to generate estimates of the statistical parameters of interest (such as the mean concentration over some area or time frame, the slope of a trend line, or some measure of variability) that address the study questions. Uncertainty in the inputs and the natural variation in sample collection are addressed in Step 5 and have a direct affect on performing Step 7.

Activities

- For all studies, specify how data will be evaluated and used to address the study question. Include a description of any statistical parameters (such as a mean, upper percentile, upper tolerance limit) to be calculated using the data. Consider use of parameters with appropriate statistical properties.

- Construct example graphics that show the intended techniques for exploring and presenting the data (e.g., cross correlation plots, box plots, spatial plots, tables). Even for descriptive studies, this exercise will help communicate the plan to all interested parties on the extracting information from the data.

- Discuss how data will be prepared for analysis, including how values below the detection limit will be handled, how outliers will be identified and handled, and any manipulations that will be performed (such as normalizing results, or calculating indicators) prior to using the data to address the study questions.

- If appropriate, construct detailed diagrams (e.g., flowcharts) that show how the various technical inputs (equations, models, data, expert opinion, regulatory requirements, etc.) will be combined to answer the study questions.

- Specify the general form of any deterministic, probabilistic, or statistical models that are needed.

- Decide on the strategy that will be used to evaluate the uncertainty or variability of the unknown parameters. Consider calculation of confidence intervals, prediction intervals, and other quantitative indicators of uncertainty. Specify how information on uncertainty will be used to evaluate the performance of the study. This could include a sensitivity analysis on a model or calculation of a specified confidence interval around a mean, correlation plot, etc.

- Develop initial ideas on how the results of the study may have potential for secondary purposes (i.e., for purposes other than those specified for the study).

Outputs

- A description of the data analysis approach.

- A description of any statistical models that will be employed.

- Documentation of how uncertainty and variability will be characterized and incorporated into the models.

- A description of how the results will be communicated to interested parties.

- Sample figures showing the intended data presentation.

- Preliminary consideration of secondary uses of the data.

3.7 STEP 7: OPTIMIZE THE DESIGN FOR COLLECTING INFORMATION

Sampling is the selection and collection of physical specimens from units of the population and the measurement of attributes on those specimens that are directly relevant to the study objective. The goal of this step is to optimize the design for data collection and this step may mean either (a) the most resource-effective data collection program that is sufficient to fulfill the study objective or (b) the data collection program that maximizes the amount of information available for synthesis and analysis within a fixed budget.

Samples may be taken from the target population in either of two ways; probability-based and judgmental. In a probability-based sampling scheme, each sampling unit has a known probability of being selected, and only those sampling units selected will be observed or measured to provide the data for the study. In a judgmental sampling scheme there is no "known probability" of being selected, as the samples are chosen or obtained only by discretion of the person in charge of the sampling effort. The difference is crucial to the inference drawn from these samples. For specific guidance on sampling designs, refer to *Guidance on Choosing a Sampling Design for Environmental Data Collection (EPA QA/G-5S)* (EPA, 2002b).

To support a statistical design, it is necessary to obtain a relevant estimate of the variance associated with the parameter of interest, determine what statistical procedure will be used, and analyzing the information quality constraints generated from Step 5 of the PAC Process. Obtaining the relevant estimate of variance is critically important and greatly affects the ability of the design team to determine the expected performance of different design alternatives. In some cases it will be advisable to design and implement a preliminary or pilot study to obtain this information and determine the relative contributions of spatial and measurement variability. In other cases, historical data will be adequate to inform the design.

For studies that involve the design and implementation of an experiment, techniques for experimental design will be involved. Technical guidance on experimental design is beyond the scope of this guide; however, it is appropriate to develop performance and acceptance criteria for studies that involve experiments. Optimization in this case may involve varying the number of sampling units, number of treatments, number of replicates per treatment, allocation of treatments to sampling units, methods for obtaining observations related to the treatments, and methods for ensuring randomness. For specific information on experimental design, refer to Box et al. (1978), Cochran (1977), or other texts on the subject of experimental design.

Activities

- Obtain relevant estimates of variance to support the design process. If necessary, design and conduct a preliminary study to generate the needed information and so use the PAC Process in an iterative fashion.

- Define the statistical approach that will be used to derive information from data.

- Determine the amount of data required to fulfill the study objective and performance and acceptance criteria.

- Evaluate various data collection alternatives and the cost of each.

- Define the time period for collection of data.

Outputs

- A data collection plan (QA Project Plan, or Sampling & Analysis Plan) that best meets the objective of either (a) providing sufficient data of adequate quality at the least cost or (b) maximizing information quality within a fixed budget.

CHAPTER 4

ILLUSTRATIVE EXAMPLES OF PERFORMANCE AND ACCEPTANCE CRITERIA

This chapter shows performance and acceptance criteria that were developed for a variety of environmental information collection activities. The examples are intentionally brief.

This chapter provides several relatively brief examples of the outputs from systematic planning, specifically focusing on the concepts of performance and acceptance criteria (i.e., Steps 5–7). More extensive case studies, discussing rationale and planning activities as well as the outputs, are provided in Chapter 5.

The process described in this document is EPA's recommended systematic planning approach for a wide variety of environmental data collection activities. While it is not possible to illustrate all of the applicable types of projects, a sampling of some of the activities that may benefit from this guidance include the following:

- *Surveys:* What are the distributions of direct and indirect water ingestion for specified sub-populations in the U.S. as well as the general U.S. population?

- *Risk assessment studies:* What are the total human environmental exposures to metals, pesticides, and volatile organic compounds in a specified area?

- *Laboratory experiments:* How can we better simulate the human digestive processes so that we can improve our estimates of the bioavailability of soil contaminants?

- *Environmental modeling:* What are the best methods for using field monitoring data to validate a computer-based modeling system for estimating air pollution?

- *Monitoring for pollution:* What sampling plan will best evaluate the presence of bacteria at recreational beaches?

- *Ecological population studies:* Do extremely low frequency (ELF) communications systems and their associated electromagnetic fields (EMFs) cause changes in plant and animal populations?

- *Process control studies:* What are the optimal control settings for a biomass gasification plant to maintain efficient and effective operation?

- *Demonstration projects:* How effective is a proposed new technology in remediating volatile organic compounds in soils?

The following subsections provide examples of performance and acceptance criteria for several illustrative environmental information collection activities. Each example is introduced with a brief background discussion, setting the stage for the study, and then illustrative performance and acceptance criteria (i.e., the outputs from Step 5, 6, and/or 7) are listed.

4.1 LABORATORY METHOD DEVELOPMENT PROJECT

Suppose we are conducting a laboratory method development project for the analysis of dioxins in cow's milk because no such validated method currently exists. As a first step, we are considering three potential methods which have been used for dioxins or similar compounds in other matrices. To pass our initial screen and be considered for further evaluation, a method must be reasonably free from bias. Since no decisions are being made with direct, serious health or environmental consequences, it is sufficient that only the information type and quality must be formally specified. Also, since this study will be generating new quantitative data, the design is developed in terms of performance, rather than acceptance, criteria.

Information Quality (Step 5)

Performance criterion: Percent recovery of chemical concentrations in spiked cow's milk samples (as measured by specific dioxin congeners) between 50% and 150%.

4.2 STUDY OF ANNUAL AIR POLLUTION TRENDS

Several air pollution regulations have been implemented at the federal, state, and local levels, many of which require routine monitoring of air quality. The objective of this study is to document the annual trends of various air pollutant concentrations within a given geographic area over the past 10 years. Since we plan to use existing data collected by many different organizations, the type of available information, as well as the quality and quantity of data, should be considered in setting acceptance criteria. The EPA has established a single major database, Aerometric Information Retrieval System (AIRS), as a means of merging most of the air monitoring data collected across the U.S. While not every U.S. monitoring program reports to AIRS, those that do participate agree to make their data available to the EPA and public after it has been cleaned and submitted in a standardized format. In addition, within AIRS, the frequency at which the data were collected and the technology used to measure the pollutants are noted.

Information Quality (Step 5)

Acceptance criterion 1: Data from the monitoring program must be specifically for our geographic area of interest, with a minimum of 5 years of weekly data, and measurements from at least 3 days each week pooled into the weekly average.

Acceptance criterion 2: Monitoring equipment must have sufficient sensitivity so that at least 75% of the pollutant concentration data are above the detection limit.

Acceptance criterion 3: To help interpret annual changes versus seasonal variations, weekly data are required from at least 80% of the weeks within each season and year covered by the monitoring program.

4.3 EVALUATION OF CAPILLARY VERSUS VENOUS BLOOD-LEAD TESTING

Despite the known potential for bias, several lead poisoning prevention programs and doctors use capillary testing (finger pricking), as opposed to the more accurate venous testing, in order to measure the blood-lead (PbB) levels of children. In order to confidently utilize blood-lead measurements of both types, an evaluation is required of the extent of bias in the capillary data. If possible, we would like to conduct this evaluation using existing data from a standardized database maintained by the Centers for Disease Control and Prevention. Thus the study design would incorporate acceptance criteria.

<u>Information Quality (Step 5)</u>

Acceptance criterion 1: Only data from the Centers for Disease Control and Prevention databases for those blood-lead screening programs that have a documented and effective QA program will be used.

Acceptance criterion 2: Data sets will not be used unless the analytical methods employed achieved the appropriate method detection limit.

Acceptance criterion 3: Two test measurements are required from each child (either cap/cap, ven/ven, or cap/ven), and we will use the two most recent results which were taken no more than four weeks apart.

<u>Information Synthesis (Step 6)</u>

Acceptance criterion: Errors-in-variables and random-effects regression models will be used to model the bias in capillary PbB relative to venous PbB statistically.

4.4 FIELD-BASED MEASUREMENT METHOD FOR MERCURY CONTAMINATED SOIL

An EPA Region is considering adopting a new field-based measurement method for the characterization of soil mercury levels at contaminated sites. This method promises to produce real-time results at a considerable savings compared to fixed-laboratory analytical methods. To evaluate the performance of the new method, the region will employ it at a site that has recently been characterized using conventional methods. The principal study goal is to determine whether the field-based method achieves satisfactory performance in determining concentrations of mercuric chloride in soil compared to fixed-laboratory techniques. Because the project is to collect new data, the design will incorporate performance criteria.

<u>Information Quality (Step 5)</u>

Performance criterion 1: Analytical accuracy ±10% and precision ±20%.

Performance criterion 2: Analytical limit of detection of 5 mg mercuric chloride per kilogram soil.

<u>Information Synthesis (Step 6)</u>

Performance criterion: This data analysis will involve estimation of the geometric mean mercury concentration.

<u>Optimize Design (Step 7)</u>

Performance criterion: N = 41 soil samples, derived as the number of samples that could be analyzed for 75% of the cost of the previous fixed-laboratory characterization.

4.5 PROCESS CONTROL STUDY AT A BIOMASS GASIFICATION PILOT PLANT

The purpose of the project is to verify design and operating characteristics of a low-pressure, indirectly-heated gasifier technology. The gasifier operates by heating biomass (wood chips, whole tree chips, municipal solid waste, etc.) with circulating hot sand until the wood breaks apart into its constituent chemical components and the carbon, hydrogen, and oxygen in the biomass form combustible gases. The result is a clean-burning gas with a medium heat content that can fuel commercial gas turbines. This project involves long-term operation and testing, including trials of the entire system that will establish the most important parameters for system reliability and characterize system operation and performance. Because it is concerned with new tests and data, this project design will incorporate performance criteria.

<u>Information Quality (Step 5)</u>

Performance criterion 1: Changes in process input and control parameters of 10% or greater will be considered outside the range of normal variation, and thus considered to be a shift of conditions that should initiate a change in output.

Performance criterion 2: The low-pressure gasifier is expected to have a 122 Mwe net system output when operating in a steady state condition. Normal hourly fluctuation around this value is ±20 Mwe. Any deviations from this value would be attributable to perturbations of the inputs or changes in operating conditions.

<u>Information Synthesis (Step 6)</u>

Performance criterion: Each of the components in the system will be monitored for deviation from its original calibrated condition. Shifts from the calibrated condition will be monitored using control charts and trend lines. Natural variation when the system is in steady state and being operated to maintain a target range of biogas output will be assessed using repeated samples over time to calculate short-term variation (i.e., “noise”). When the system is being operated in experimental mode (i.e., changing the inputs or “signal”), the researchers are searching for optimal levels and improved performance.

4.6 TECHNOLOGY DEMONSTRATION FOR REMOVAL OF SOIL CONTAMINANTS

Environmental regulators and a site owner are concerned about effectiveness and costs associated with remediation of a site that has diesel and gasoline fuel spills in the soil. A new technology uses standard soil vapor extraction in combination with *in situ* heating of soil in order to strip off volatile organics and greatly reduce the time needed for remediation relative to unaided soil vapor extraction. This new approach, although it has shown some promise, has not been demonstrated thoroughly enough to allow for its proposal and/or acceptance as a preferred method for removing volatile organics from contaminated soils. Both the regulators and the site owner are interested in defining the capabilities of the *in situ* heating of soil in stripping of volatile organics at a site that has diesel and gasoline fuel spills in the soil. Their main questions are the following:

1. What is the rate of contaminant extraction in each lithological unit?

2. What range of fuel components are removed in each lithological unit?

3. What percentage of each fuel component is removed from each lithological unit?

As the project relies on new data, it is designed in terms of performance criteria.

<u>Information Synthesis (Step 6)</u>

Performance criterion: For the first question, the monitoring data from each lithological unit over time for each volatile constituent will be fit to a mathematical model and model parameter uncertainties estimated. It is expected that the extraction rate will vary over time. For the second and third questions, the pre-extraction mass estimate and post-extraction mass estimate will be calculated as averages under the assumption that the samples are far enough from each other to be independent. For the second question in particular, the difference between the pre-extraction mass estimate and the post-extraction mass estimate will be the measure. Those constituents whose estimated removal is significantly different from 0 will be considered as treatable by the *in situ* heating technology. For the last question, the ratio of the post-extraction mass estimate to the pre-extraction mass estimate will serve as the estimator, and uncertainties will be developed by propagation of error techniques.

CHAPTER 5

CASE STUDIES OF SYSTEMATIC PLANNING USING THE PAC PROCESS

This chapter contains six case studies showing the use of the PAC Process to generate performance and acceptance criteria. These studies are not intended to be prescriptive but to illustrate how the investigators in this particular instance arrived at their criteria.

The following six case studies are provided as examples of the seven-step PAC Process as it applies to various research or investigatory studies. The case studies were chosen to represent a range of increasingly challenging environmental data collection and analysis projects:

Case Study 1 is a laboratory study collecting new data.

Case Study 2 is a field study using existing monitoring data.

Case Study 3 is a field study collecting new data.

Case Study 4 is a modeling study utilizing existing secondary information.

Case Study 5 is a field study that will collect new, but preliminary, data that will be used for more definitive monitoring plans.

Case Study 6 is a field study to collect new data with the performance criteria specified quantitatively and the information quantity determined statistically.

These examples are based on actual studies and projects, although specific details have been altered to prevent erroneous comparison with ongoing studies and projects.

5.1 LABORATORY STUDY OF A PHYSIOLOGICALLY-BASED EXTRACTION PROCEDURE

This example is for a new data collection effort where the design will consider performance criteria. This study involves a laboratory research effort whose primary focus is obtaining reliable methods development information. The study planners have chosen to emphasize Step 5, which they specify quantitatively, but the last two steps are handled more qualitatively due to limitation of resources.

Step 1: State the Problem

Background

Soil contaminants pose a threat to humans through several exposure pathways, including dermal contact, inhalation of contaminated vapors or particulate material, and ingestion of contaminated soils. Remedial technologies have been developed to reduce these risks of exposure, and their performance has been evaluated based on the total mass of extractable contaminants from treated soils. In cases where soil ingestion is the primary exposure pathway of concern, this method of evaluation may not properly estimate the actual availability of contaminants to humans. To better estimate exposure due to soil ingestion, a physiologically-based extraction procedure (PBEP) has been developed. The PBEP is an in vitro extraction procedure designed to simulate human digestive processes and to evaluate contaminant desorption under physiological conditions. The PBEP contains several modules, each representing a different organ in the human gastrointestinal tract.

Output

The stomach module of the PBEP has been developed and is now in the testing phase. Briefly, the protocol for the stomach module can be summarized as follows:

1. Prepare soil and stomach fluids for use in the experiment.
2. Create samples by filling bottles with soil and mixing the appropriate fluid.
3. Incubate samples while shaking at the specified speed and length of time.
4. Separate the liquid and solid portions of the sample and analyze both for the concentrations of individual polycyclic aromatic hydrocarbons (PAHs).

There are several operating parameters in the protocol that can be varied. These parameters include fluid composition, fluid pH, incubation time, and shaker speed. The relationship between various levels of these parameters and PAH desorption has yet to be determined. A series of studies has been proposed to examine the effects of the operating parameters on PAH desorption in the PBEP. In this particular study, three of the operating parameters will be examined: fluid type, presence of food, and pH.

There are several constraints that will affect the design of the experiment. First, the shaker/incubator apparatus allows for only four samples (plus one blank and one spiked sample) to be run per batch. Second, only one batch per day can be run. Third, because of the degradation characteristics of the fluids, it is preferable to perform all the experiments using one fluid before performing the experiments using another fluid. Furthermore, no more than eight batches can be run from a single fluid before degradation begins. Finally, cost, time, and replication constraints limit the total number of samples to 50. Furthermore, replication of each treatment combination is required.

Step 2: Identify the Study Questions

> **Background**
> *There are several factors that might influence the desorption of PAHs in the PBEP stomach module. Trying to evaluate all the factors at once will be costly and time-consuming. As a result, this preliminary study will focus on only a few of the factors, and the questions of interest will address those factors. Results from this study can then be used in designing further studies to examine the PBEP stomach module.*

Output

As noted above, the primary goal of the current experiment is to evaluate the effects of fluid type, food, and incubation time on contaminant desorption in the stomach step of the PBEP. The specific questions of interest include the following:

1. Does the presence of food in the fluid inhibit PAH desorption?
2. Does pH level vary depending upon the pH level of the fluid?
3. Does PAH desorption differ for different types of fluid?
4. How do food, fluid, and pH interact with respect to PAH desorption?

The objective of this study is not to provide definitive information about the relationship between PAH desorption and the operating parameters, but rather to determine which parameters may be related and should be examined more closely in subsequent studies.

Step 3: Identify Types of Information Needed

> **Background**
> *In this laboratory study, the operating parameters (fluid type and pH, presence or absence of food) are experimentally controlled and can be considered constant and known. Results from chemical analyses of the samples for PAH levels will vary and need to be controlled by specifying appropriate MQOs.*

Output

Quantitative data on fluid type, amount of food, and pH level will be measured, along with resulting PAH desorption concentrations. Other operating parameters such as incubation time and shaker speed are not of interest in this study and will be held constant for all the samples. However, the shaker speed and incubation time will be recorded for each batch. In addition to the operating parameters, because there may be differences in PAH desorption related to the batch, the batch number for each of the samples will also need to be known. Analysis of the sample after incubation includes chemical analysis of the liquid and solid portions of the sample. Both portions of the sample will be analyzed for PAH concentrations. Concentrations will be measured for individual PAHs, although the concentrations will be combined with each ring class and over all ring classes

Step 4: Establish Study Design Constraints

> **Background**
> *Technically, the population of interest is the set of all experiments that can be run by varying the fluid type; its pH; and the presence or absence of food with a fixed soil, incubation time, and shaker speed. However, here the population will be restricted to only those fluid types that fit within the design constraints discussed in Step 1.*

Output

The characteristics that define the population of interest are the PBEP operating parameters. Specifically, the population that will be examined in this study is the population of all samples that can be created for each fluid type and pH level either without food or with a specified amount of food present. The population will be further restricted to all such samples that are incubated for a specified amount of time at a specified shaker speed. Due to cost and time restrictions, the experiment will be performed using only two fluid types, which have been chosen as most representative from among several different types of fluids that have been created to simulate stomach fluid. The number of pH levels will also be restricted.

Step 5: Specify Information Quality

> **Background**
> *Information about the fluid type, its pH level, and whether food is present in the sample, are the factors of interest in the study. Laboratory analysis for PAH levels in solid and liquid samples provides information about a large number of individual PAHs. Examining the effects of this information for each individual PAH can be done with existing statistical software but provide results that are too detailed to produce an overall view of the effects of the factors. Providing results for individual ring classes may provide a better way of summarization.*

Output

Performance Criterion 1, Precision: If the difference between each replicate exceeds 15%, then flag the data point for future analyses and investigate the reason.

Performance Criterion 2, Bias: If the results of the blank and spiked samples supply data indicate the existence of bias, then flag the data and investigate further.

Performance Criterion 3, Sensitivity: Since these experiments investigate PBEP, the prepared samples should not pose any problems in terms of limit of detection or calibration. However, if a sample is tested outside the range of detection, then flag the data point for future analyses and investigate the reason.

Performance Criterion 4, Completeness: There must be at least two valid measurements (precise, lacking bias, and within range of detection) per combination in order to continue the investigation.

Step 6: Develop a Strategy for Information Synthesis

> **Background**
> *The analysis of data must be based on a variable that clearly measures the ability of the PBEP experiment to desorb PAHs. Such a variable must be based on the amount of PAH in the original soil sample, the amount in the solid portion of the sample after incubation, and the amount in the liquid portion of the sample after incubation. Evaluation of the effects of the factors of interest in this study will require statistical analysis-of-variance or a similar procedure. Because the PAH desorption will be measured for each ring class, the response variable is multivariate in nature, so a multivariate analysis should be considered unless the data indicate that it is not necessary. The analysis should also address the assumptions underlying the analysis-of-variance in order to determine whether it is appropriate for the data.*

Output

Performance Criteria: The variable upon which the analysis will be based is the percentage of PAH desorbed into the solution. Desorption recoveries will be calculated for each individual PAH, for all PAHs within each ring class, and for all PAHs. Multivariate and univariate analysis-of-variance (ANOVA) methods will be used to determine whether fluid pH and presence of food influence the desorption recoveries of PAHs. The response variables in the multivariate ANOVA (MANOVA) will be desorption recoveries for 2-ring, 3-ring, 4-ring, 5-ring, and 6-ring PAHs. The response variable in the ANOVA procedure will be the desorption recovery for total PAHs. When significant effects are identified, Tukey multiple comparisons will be used to identify which combinations of pH and food conditions yield the highest desorption recoveries. Prior to these analyses, the data will be examined to determine whether any transformation of the data is needed to adhere to the assumptions of the ANOVA procedures. In addition, a correlation analysis will be used to determine whether PAH desorption recoveries among the ring classes are independent.

Step 7: Optimize the Design for Collecting Information

Background
The design of the study will be driven by laboratory and cost constraints. There will be two levels of the factor representing the presence of food in the sample. The preferred number of fluid pH levels is four. Thus, there are eight combinations of these two factors. The limit of four samples per batch imposes a blocking effect based on the incubation batch that will need to be incorporated into the statistical model. There is also a strong preference for several replicate samples for each combination of factors. With eight factor combinations per fluid and a limit of 50 samples, either two replicates can be done for each of three fluids or three replicates can be done using two fluids. This will result in 48 samples to be divided among 12 batches. An appropriate design needs to be found that meets these criteria.

Output

Performance Criteria: Due to budget and laboratory constraints (Step 1), 48 samples will be analyzed, considering two fluid types, four pH levels, and the presence or absence of food.

The study that is proposed examines two fluid types and four levels of pH. There are naturally two levels of the factor relating to the presence of food in the sample (present, absent). The resulting set of 16 treatment combinations can be replicated three times, for a total of 48 samples. A design that would require three replicates has been derived using experimental design theory and is outlined in Table 2. In this design, the eight treatment combinations within each fluid type are divided into two groups of four, each group corresponding to a single incubation batch. A pair of batches will contain each of the eight treatment combinations, so that each pair of batches constitutes one replicate.

Table 2. Proposed Experimental Design for PBEP

Replicate	Batch (Block)	Treatment Combinations*			
		1	2	3	4
1	1	A1	A2	P3	P4
	2	P1	P2	A3	A4
2	3	A1	P2	A3	P4
	4	P1	A2	P3	A4
3	5	A1	P2	P3	A4
	6	P1	A2	A3	P4

*The first character of each sample indicates the presence (P) or absence (A) of food and the second character represents the pH level (1=lowest to 4=highest).

For each fluid, there will be three replicates as shown in Table 2. Thus, the full experimental design consists of three replicates (each containing two batches of four samples) for each fluid, for a total of 48 samples. Table 2 does not show the precise order in which the experiments will be run, randomization in the order of batches and replicates within batches will be necessary.

5.2 EVALUATION OF POTENTIAL INDICATOR PARAMETERS FOR METALS CONTAMINATION IN SURFACE WATER

This example describes a study using existing surface water monitoring data. The plan is to develop estimates that may help to streamline a surface water management program. Acceptance criteria are developed to circumscribe the data set to be used in the study.

Step 1: State the Problem

Background
Surface water monitoring has been ongoing at Deep River for about a decade. It is anticipated that existing data will be adequate for purposes of this study. However, since data have been collected and analyzed using a variety of methods, it will be important to specify acceptance criteria that assure data comparability.

Output

Previous studies at other facilities have suggested a relationship between total suspended solids (TSS) and the levels of certain contaminants in surface water. If it is confirmed that this correlation pertains to Deep River, then it may be feasible to use TSS as an indicator parameter in the design basis of future pond operations. Furthermore, since TSS levels can be determined only through laboratory analysis, it is also worthwhile to evaluate potential real-time indicators such as turbidity.

Step 2: Identify the Study Questions

Background
Deep River is a federal facility where past operations may have contributed to the release of contaminants, including various metals. Surface water features at Deep River include a system of perennial and intermittent stream channels that drain into a major river basin and a series of detention ponds designed to manage process effluents and storm water. As part of ongoing efforts to enhance the operation of the detention ponds, facility managers wish to assess whether parameters such as TSS can be used as indicators of metals concentrations. They have decided to conduct a special study aimed at addressing this issue.

Output

The study will address the following questions:

1. Is there is a quantifiable correlation between metals concentrations and TSS levels in samples taken from Deep River surface water?

2. Is there a similar correlation between metals concentrations and real-time turbidity measurements?

The metals to be included in the study will be cadmium, chromium, and zinc. If satisfactory correlations are established, then a future phase of this project will assess how the correlations can be used as an input to future detention pond design and operational specifications.

Step 3: Identify Types of Information Needed

> **Background**
> *Sampling stations for collection of surface water samples in streams and detention ponds have been in place at Deep River for a minimum of three years and a maximum of 10. All analytical methods used are EPA-approved methods found in 40 CFR Part 136.*

Output

The information needs for this study include the following:

1. Concentrations of cadmium, chromium, and zinc in surface water samples collected at specified Deep River sampling stations and analyzed using EPA Method 200.7 (for chromium and zinc) and 200.8 (for cadmium).

2. Measurements of TSS at the same stations using EPA Method 160.2.

3. Measurements of turbidity at the same stations using real-time continuous monitoring probes as specified in the *Deep River Surface Water Monitoring Plan*.

4. Estimates of flow rates at each station at the time of sample collection.

Step 4: Establish Study Design Constraints

> **Background**
> *For the last two years, surface water sampling has been conducted based on a documented monitoring plan that includes a full QA program. Two analytical laboratories support the program; for the past year they have used identical analytical methods, but prior to that there were some variations, including different sample preparation procedures preceding TSS measurements. For this study, it is important to select data meeting pertinent measurement quality objectives and analyzed with identical methods.*

Output

Eight monitoring locations have been identified for this study: three at upstream locations in the vicinity of past or present industrial operations; three in downstream stream channels; and two at detention pond influent points. The upstream locations were selected because they would provide a minimum of two hours warning before the same water arrived at the detention ponds. Data will be selected from a data set consisting of all relevant measurements conducted at the identified stations within a continuous 12-month period during which all identified stations remained in operation.

Step 5: Specify Information Quality

> **Background**
> *For TSS or a related parameter to function as an indicator of metals contamination, it is important to establish a spatio-temporal relationship between TSS and metals levels. In other words, we are most interested in assessing whether high upstream (e.g., in an industrial area) TSS levels at a given point in time are correlated with high downstream (e.g., at a detention pond) levels in metals at a later point in time. It is this predictive capability that would be useful for detention pond management.*

Output

Acceptance Criterion 1, Representativeness: Data will be chosen to represent the full range of values for each measured parameter and the full range of flow conditions experienced during that 12-month period. To the extent possible, data will also be chosen to represent upstream-downstream temporal succession—i.e., downstream sampling events on the next day (or later the same day) following upstream sampling events.

Acceptance Criterion 2, Sensitivity: Because it is important to determine whether correlations hold at low parameter values as well as high ones, laboratory data selected for this study should have achieved the following detection limits:

Cadmium	1 µg/L	Zinc	20 µg/L
Chromium	5 µg/L	TSS	3 mg/L

Acceptance Criterion 3, Comparability: All water samples used in this study should have been collected using procedures specified in the *Deep River Surface Water Monitoring Plan*. All laboratories analyses should have been conducted using the EPA-approved analytical methods identified in Step 3 above, including uniform sample preparation for TSS analyses.

Step 6: Develop a Strategy for Information Synthesis

Background
The size of the study data set was chosen based on the input of the statistician who will perform the calculations. The work group conducted a preliminary data screening to determine whether these acceptance criteria would be achievable. If that had proved untrue, the only alternative might have been to undertake a new sampling program.

Output

Acceptance Criteria: Study data will first be evaluated by means of exploratory data analysis, including summary statistics such as mean, median, variance, standard deviation, and range. R^2 values will be calculated to evaluate the relationship between concentrations of each metal and TSS as well as the relationship between metals data and turbidity values. Comparisons will be made for each individual sampling station and event as well as among the three sectors (upstream, downstream, pond intake). Temporal plot lines will be displayed graphically to evaluate trends, for instance between upstream TSS levels at time #1 and downstream metals values at time #2.

Step 7: Optimize the Design for Collecting Information

Background
A variety of statistical displays and calculations will be used to assess the study results and frame them in a fashion that will support the work group during the next project phase. The work group's assumption at this point is that an 80% correlation will be adequate for continuing with the project.

Output

Acceptance Criteria: For each sampling station, a sampling event that produces data for the 3 metals, TSS, and turbidity will be termed a "data cluster" for purposes of this study. For each station, this study will incorporate a total of 20 data clusters that were collected within the 12-month study period. For the downstream and pond intake sampling stations, at least 10 of those 20 clusters will represent “successional” sampling events, i.e., sampling conducted no more than 24 hours later than corresponding upstream sampling events.

5.3 ECOLOGICAL STUDY OF BIRD POPULATION SIZE

This case study describes a new data collection effort, with associated performance criteria that are specified qualitatively and quantitatively. Potential health risk from electromagnetic fields remains a controversial issue with several research questions unanswered. Therefore, this study is considered a research project.

Step 1: State the Problem

Background
Eleven studies aimed to measure the significance of possible environmental effects caused by ELF EMFs. The organisms and ecological relationships that were selected for the various studies were chosen primarily for two reasons: they seemed relatively important to the ecosystem and they appeared as good representatives for large taxonomic groups that have shown ELF EMF effects in the past (lab or field experiments). Some subjects were also chosen based on the concerns of local residents. As a result of the wide variety of testing subjects, the research teams addressed a broad range of testing questions. Though each question ultimately pertained to the potential impact of the EMFs, the ways in which specific organisms, general species, and ecological relationships may change due to electromagnetic fields may differ. One of the studies examined the local bird population in the area of the ELF facility. The remainder of this case study addresses whether or not the bird population size is affected by the ELF communication system.

Output

Controversy still exists concerning the potential impact that ELF EMFs may actually have on the environment. Scientific theories do not corroborate how the ecosystem surrounding the communication facilities may change due to exposure to EMFs. Theories suggest physiological, developmental, and/or behavioral changes in individual organisms or their communities. Some scientists believe biological responses attributed to EMF exposures have been reproducibly demonstrated, others are skeptical about the documentation of such responses, and yet others have expressed the opinion that such responses violate fundamental laws of physics and therefore are physically and biologically impossible. Hence, the U.S. Navy contracted several researchers to begin studying the problem to discover if and how the ELF system impacts the local environment of the transmitting facilities.

Step 2: Identify the Study Questions

Background
Since 1950 the U.S. Navy has been exploring the use of an ELF communications system in order to transmit messages to submarines located anywhere in the world. In 1969, the first of two transmitting facilities was built as an experimental station. Since then, the facility was upgraded, and in 1985, it became fully operational. However, a potential problem with the facility pertains to environmental exposure to EMFs generated by the ELF communication system. In order to investigate the possible EMF hazard, a research institute agreed to provide management and scientific support for establishing an ecological monitoring program.

Output

The primary goal for the research project is to answer the following question:

> Does the number of birds, in general, or within selected guilds[1], differ between areas close to the ELF antenna versus control areas presumed to be far enough away to be unaffected by the antenna?

Since birds use the earth's magnetic field for orientation during migration, they are important organisms to consider in an assessment of EMF impacts. Though several researchers have studied the effects of ELF EMFs on most aspects of a bird species' life history, the subject is still poorly understood. In part, this is due to the inability of past study designs to differentiate the effect of the actual ELF EMFs from the effect of alteration in habitat caused by building the facility and installing the antennas. Hence, previous researchers have investigated the combined effect of habitat alteration and EMFs. In contrast to previous research, this study hopes to isolate and investigate the effects of ELF EMFs on local bird species and communities.

Step 3: Identify Types of Information Needed

Background
There are several factors that must be taken into consideration when attempting to pinpoint the effects of EMFs. Therefore, not only do the researchers need information pertaining to the number of birds within a specified region, they must also obtain information for classifying bird guilds, habitat, right-of-way (ROW) effects, and EMF intensity levels. These latter factors will provide explanatory information to help account for possible confounding effects in bird population sizes.

[1]A documented classification of birds that pertains to nesting site, food, breeding habits and migration patterns

Output

This data collection effort will concentrate on quantitative measures related to the bird populations, as well as EMF intensities in the area of ELF transmitters and in background areas.

Number of Birds. Bird population sizes are difficult to measure and vary significantly due to natural forces such as time of year, weather, competitor populations, disease, etc. In addition, ELF EMF intensity varies over time, leading to differences over time in bird exposures to EMFs. Therefore, multiple counts of the population over time should be taken. The data should be collected during optimal times of the day and year when the birds are most visible and active.

Guilds. The primary question for this study has two parts. The first question is broad, "Does the bird population size change?" The second, however, narrows the question and asks, "Does the population size change within selected guilds?" A guild classifies the birds into groups based on feeding strategies, specific nesting areas, breeding habits, and migratory patterns. Therefore, the researchers need specific information about each indigenous species in order to accurately classify the birds.

Habitat. Areas of similar vegetation should also have similar bird communities. Hence, consistent, detailed notes about the habitat are necessary for the study. Densities of trees, shrubs, forbs, and garminoids will be measured. Control transects (considered unaffected by the ELF transmitters) will be paired with treatment transects by similar habitats.

Right-of-Way (ROW). The "right-of-way" defines that portion of the forest which was cleared to install the ELF antenna. Beyond the effects of ELF EMFs, the act of clearing the forest may change the bird population. Thus, controlling for ROW in the study design is important to isolate EMF effects from ROW effects.

EMF Intensity Levels. The EMFs produced by the ELF system are not consistent throughout the year, nor from site to site. At various times in the day, different portions of the antennas are turned on or off, with varying modulations, frequencies, current intensities, and phase angles. In general, however, when the antennas are on, the EMF frequency modulates between 72 and 80 Hz depending upon the message that is being transmitted. In addition, the farther the test site is from the antenna, the lower the levels of EMF intensity, with intensity decreasing in proportion to the square of the distance from the antenna.

Ambient Intensity Levels. Any source of EMFs, whether it be an ELF communication system or a typical power line, creates fields that span over broad areas. In fact, selecting a control site that is never exposed to EMFs is practically impossible. Thus, adjustments must be made to the study so that (1) the difference between the intensity levels of EMFs between treatment and control sites is consistent, and (2) any interferences of non-ELF EMFs are taken into consideration so that the treatment effects are not masked. In general, both the treatment and control sites are expected to be exposed to 60 Hz fields (most U.S. equipment produces EMFs at 60 Hz).

Step 4: Establish Study Design Constraints

> **Background**
> *The target population for this effort is all birds surrounding the ELF antennas; however, various sampling biases are possible. Therefore, we must consider the sampling method as well as the population of interest in order to make a proper interpretation of the project results.*

Output

The target population is all non-endangered, observable birds that are within a reasonable range of the transmitting facilities and are potentially affected by the ELF communications system. Where, when, and how a data collector looks for birds could limit the target population. For this effort, the data are collected within 500-meter segments of five transects within either treatment or control sites. The observers are trained to spot birds by sight and sound that were within 100 meters of the transect center line. Beginning ½ hour before and continuing until 4.5 hours after sunrise, the observers walk at a pace of one km/hour and note all of the birds that they hear or see.

Birds that are not documented fall into one of the following four categories:

1. The bird appeared farther than 100 meters away from the center line.
2. The bird was flying above the canopy line.
3. The bird was considered rare or potentially endangered (this was outside the scope of this study).
4. The bird simply was not seen or heard.

Note that a particular bird species was not isolated for this study. However, the study was limited to the species that live within the northwest region of the U.S. and within the habitats that surround the transmitting station.

Step 5: Specify Information Quality

> **Background**
> *In terms of data collection, this effort will be rather conventional; however, importance still rests on the quality of the data. Therefore, established techniques and instruments will be used to count birds and document EMF intensity levels.*

Output

Performance Criteria: Quality requirements are specified for each of the primary types of information that will be collected.

Number of Birds. The approach to be used for counting birds is well established and considered "standard practice." All observers will be well qualified and experienced in the identification of birds by sight and sound, and training sessions will be conducted prior to data collection. Note that due to the nature of this study, biases may be introduced by several factors. For example, loud, active, brilliantly-colored, large birds that are typically male will probably be spotted more often than small, camouflaged birds that may spend a lot of time in the nest, such as female or young birds. However, since both the treatment and control sites will suffer from the same limitation, these biases in the data collection should not influence the results of the study. Furthermore, temporal and spatial biases will be controlled in the sampling process. For example, temporal variation in bird activity will be controlled by simultaneously collecting data from control and treatment sites by two observers. As for spatial variability, the starting points for each transect will be randomly selected, and the direction of travel from the starting points will be randomly determined as well.

Habitat and Guilds. Vegetation in all study areas will be measured over a two-year period. This time span was chosen to control for seasonal variation in vegetation growth, and the method to be used to measure vegetation has been successfully implemented in past investigations. Similarly, population guilds are well known from past ecological studies in the area over the past several years, although their specific population sizes will, of course, need to be measured for this specific study area.

Right-of-Way (ROW). The treatment transects will be designed to reduce or eliminate ROW edge effects by placing a 25-meter buffer between the ROW and the sampling areas (i.e., the center line for each treatment transect will be 125 meters away from the antenna). The effect of the ROW could extend beyond 25 meters, but to increase the buffer distance would also decrease the intensity of EMFs within the treatment transects. Therefore, a balance was found between the distance from the ROW and the diminished EMF intensity away from the antenna.

EMF Intensity Levels. The ability to measure low-level EMFs depends on, among other things, the sensitivity of the instrument. The magnetic-field probe voltage output equipment will be calibrated for measuring levels on the order of 100 mG.

Step 6: Develop a Strategy for Information Synthesis

Background

Collaboration is required between ecologists in charge of data collection and statisticians in charge of data synthesis. The data needed for the most ideal statistical analysis of the effects of ELF EMFs may not be practically feasible to obtain. Thus, the two groups must develop careful plans for analysis of the information that will be available.

Output

Performance Criteria: Statistical analysis of the study results will be performed. Observed significance levels will be used to isolate potentially significant effects for future research efforts, rather than to determine final conclusions on the effects of EMF levels on bird populations.

In order to isolate effects of the ELF communication system, treatment and control sites will be selected primarily based upon two criteria: habitat and EMF intensity level. Treatment sites will be selected to have 10 times more EMF intensity than control sites. An analysis will be conducted to examine the difference between the paired treatment and control sites, using ANOVA, controlling for season and year. Annual differences and treatment effects will be examined for the following variables: (1) number of individuals observed in a 500-m segment and (2) number of individuals in representative guild categories. Nonparametric tests may be utilized if the variables do not meet ANOVA assumptions, even after appropriate data transformations.

Step 7: Optimize the Design for Collecting Information

> **Background**
> *Several factors must be taken into consideration to create an effective experimental design, including limitations outside the technical scope of the project (e.g., budget, time, and space). With careful planning, the hope is to obtain an adequate number of unbiased samples so that sufficient quality data are obtained regardless of logistical limitations.*

Output

Performance Criteria: The sample size will be primarily driven by the available funding and schedule afforded by the U.S. Navy.

Four times a year, five treatment and five control transects will be monitored. Within each transect, data will be collected from eight 500-meter segments. As a result, 80 samples (40 treatment and 40 control) will be collected during each of the four sampling campaigns. Figure 5 displays the design of each transect and Figure 6 provides a map of all of the transects. (Note that a transect equals eight segments or samples.) The budget ($100K per year) limits the number of personnel and hence the number of transects that can be monitored. To control the effect of a single observer, each data collector will be responsible for a single transect, or group of segments. Thus, it was not feasible to investigate more than 10 transects (five treatment and five control).

Ctl = control transect
Trt = treatment transect

Large sample sizes typically yield more accurate estimates of population parameters; however, larger sample sizes mean smaller areas (length of segment) over which the investigators may collect data. The segment length, in turn, affects the amount of time needed for an observer to collect population data. Data collection is also constrained by the time of day, with birds generally being more active during the morning hours. Therefore, in terms of time, this study is limited by two factors: (1) time the observers need to collect data over a given segment length and (2) time that the observers can collect data during the optimal time of day. Therefore, the maximum number of samples per transect that can be collected, given the time constraints, is eight 500-meter segments.

5.4 EVALUATION AND IMPROVEMENT OF THE SAPM TOXICS MODEL

This case study considers an environmental modeling study that will collect secondary data on air pollution levels. These data will be used in a highly quantitative way to help calibrate an existing air quality model. The results will be used for model development rather than regulatory purposes.

Step 1: State the Problem

Background
The System for Air Pollution Modeling (SAPM) is a modified version of a standard EPA tool designed to model long-term concentrations of hazardous air pollutants over large spatial scales. The SAPM model is intended to:

- *estimate relative contributions of broad categories of emissions sources,*
- *characterize potential public health implications of air toxics, and*
- *characterize the relationship between the geographic distribution of modeled air toxics concentrations and demographic variables.*

This example is hypothetical although modeled on actual practices. SAPM, the state of Concordia, and the pollutant benzene were arbitrarily chosen.

Output

Stakeholders wish to use the SAPM modeling system to estimate benzene concentrations in air in the state of Concordia, county by county. Before estimates made by the SAPM system can be used in decision making, however, research must be done to calibrate and, if needed, improve the ability of the SAPM model to estimate the levels of benzene and other air toxics of concern.

The objective of this evaluation is to determine the accuracy of SAPM estimates by comparing model predictions with actual data from valid, independent field monitoring programs.

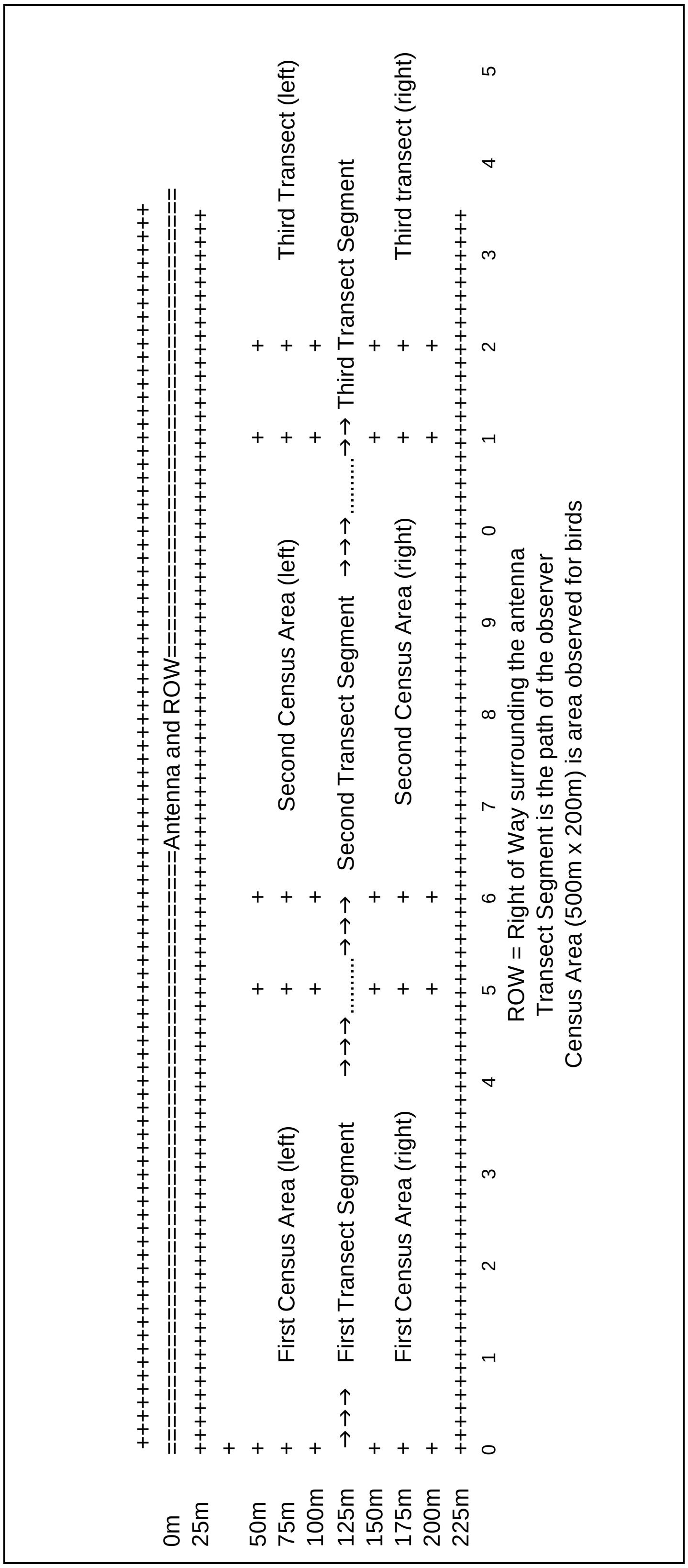

Figure 5. Map of the First Part of a Single Transect

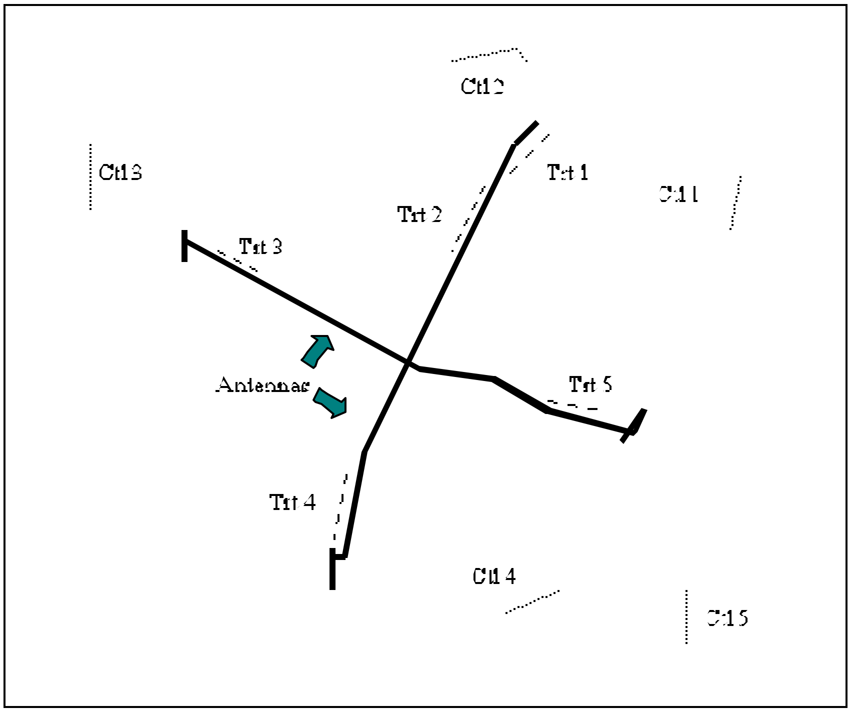

Figure 6. Map of Antennas and Placement of Transects

A planning team was selected to assess the SAPM model, including: (1) a member of the state air board familiar with ambient air monitoring programs, who will interface with other Federal and state agencies that have air monitoring data sources available; (2) a representative of the EPA who is familiar with the SAPM model and will be responsible for developing the design of the assessment; (3) a representative of a public/industry stakeholder committee who is familiar with ambient air quality; and (4) a modeler familiar with model validation techniques.

Step 2: Identify the Study Questions

> **Background**
> *The stakeholders are interested in developing an approach to assessing SAPM prediction accuracy through use of ambient air monitoring data. However, the limited budget for SAPM development and enhancement precludes the possibility of collecting new monitoring data. Therefore, this study will identify and obtain secondary data from existing air monitoring programs in order to meet the needs of SAPM development.*

Output

The main study question developed to help focus the SAPM development efforts and associated information requirements is the following:

- How accurate are the SAPM system's predictions?

Other related questions that will be dealt with during the planning process include the following:

- How should "accuracy of prediction" be defined for purposes of this evaluation?

- Within Concordia, should the evaluation be carried out separately for different strata (e.g., urban versus suburban versus rural or coastal versus mountain versus inland desert climate regions)?

- Should validation data be combined across monitoring programs and technologies?
- Which comparison databases should be used to assess the accuracy of SAPM?

- What is the comparability among the chosen databases, and what is the similarity of their results to those from SAPM?

Step 3: Identify Types of Information Needed

Background
Ideally, an evaluation of the SAPM model would directly compare the model's estimate of the annual average air toxic concentration with the true annual average in every census tract of Concordia. Unfortunately, the annual average concentrations (per HAP or other modeled compound) from actual monitoring data for each of the approximately 60,000 census tracts in the U.S. are unknown. However, air toxics monitoring data, when combined and summarized appropriately, may be used to estimate ambient concentrations for the purpose of evaluating the SAPM model. The SAPM model predicts ambient concentrations at the census tract level. Stakeholders strongly believe predictions at such a fine scale are unreliable, hence the evaluation is to be made at the county level. Therefore, SAPM census tract concentration predictions must be combined appropriately to form SAPM county-wide predictions.

The SAPM model predicts ambient concentrations at the census tract level. Stakeholders strongly believe predictions at such a fine scale are unreliable, hence the evaluation is to be made at the county level. Therefore, SAPM census tract concentration predictions must be combined appropriately to form SAPM county-wide predictions.

Output

The SAPM model predicts benzene levels by spatial location and time. Therefore, county-averaged benzene model predictions and monitoring data are required, along with specific information on the geographic location and time associated with the data.

To evaluate and improve the SAPM model estimates, information is needed from field measurements of benzene levels in ambient air, and those field measurements should be expressed in terms consistent with those produced by the SAPM model. Therefore, appropriate air monitoring data should be combined to generate an unbiased estimate of "true" county-wide annual averages.

In order to combine data, information must be obtained regarding the point-in-space (i.e., latitude and longitude) and point-in-time associated with each ambient measurement. Additionally, the geographic boundaries of counties in Concordia must be known in order to determine which data to use for combining into each county-wide average. Similarly, since SAPM predicts annual average concentrations at the census tract level, each census tract must be associated with its respective county in order to generate a county-wide SAPM prediction. Furthermore, SAPM predictions assume constant concentration levels across each census tract; therefore, each census tract's area must be obtained in order to appropriately weigh its relative contribution toward the overall county-wide average.

Step 4: Establish Study Design Constraints

Background
For this evaluation, the chosen resolution is the county level because the Concordia data available contained only one monitor per census tract but multiple monitors in some counties. This is useful for estimation of spatial measurement error for the air monitoring data. A decision error that the SAPM model is not performing adequately in a given census tract may occur if the true annual average concentration (as estimated by the local monitoring station) is significantly different from the SAPM estimate.

Output

While the planning team is ultimately interested in validating the SAPM model predictions for all chemicals and all regions of the U.S., the first phase of this assessment is targeting the average concentration of benzene at the county level in Concordia. The population of interest in this evaluation is the ambient air in Concordia, especially the average concentration of benzene in air at the surface. The parameter of interest is the annual average benzene concentration for each county.

In counties with more than one field monitoring site, those monitor measurements will be averaged. The evaluation, then, will consist of a comparison across Concordia counties of county-average SAPM predictions and county-average monitor measurements. The monitoring data used for purposes of this model evaluation are from the Concordia Air Quality Board.

Step 5: Specify Information Quality

Background
As the goal of this assessment requires the use of existing data, collected by many different organizations, the quality of the data must be considered. The pre-existing databases must have been properly verified and validated for their particular study and properly documented for secondary use.

Output

Acceptance Criterion 1: In order for a database to be acceptable for use in this model assessment, data must be oriented to the geographic area of interest (i.e., county-level data in Concordia), with a minimum of three years of weekly data and measurements from at least three days each week pooled into a weekly average.

Acceptance Criterion 2: The monitoring equipment must have sufficient sensitivity so that at least 75% of the pollutant concentration data are above the detection limit. In addition, each database must have data documented to be within ±10% accuracy and ±20% precision.

Acceptance Criterion 3: To help interpret annual changes versus seasonal variations, weekly data are required from at least 80% of the weeks within each season and year covered by the monitoring program.

Step 6: Develop a Strategy for Information Synthesis

Background
In order to ensure effective calibration of the SAPM model, the air monitoring data should be powerful enough (i.e., accurate and plentiful) to detect a bias in SAPM prediction from the true county-wide annual average benzene concentrations. If there are insufficient data or if the network is not arranged appropriately, results might be too variable to conclusively reveal a bias. These factors could be evaluated statistically to determine the amount of monitoring data required. However, since this is a model development project, the data collection will be constrained by budget considerations rather than statistical requirements.

Output

Acceptance Criteria: The model evaluation will be defined here as estimation of the slope of a statistical linear regression model where the SAPM-predicted annual average benzene concentration is expressed as a function of the annual average concentration based on monitoring measurements. Actual concentration measurements are often well represented by a log-normal distribution as they are bounded below by zero and have occasional large values, and log-concentrations may be utilized in the model.

Note that in the modeling approach, the monitoring data will be adjusted for measurement error in order to estimate the relationship between SAPM predictions and "true" annual average concentrations rather than estimated annual average concentrations. The planning team feels that this adjustment is crucial because failure to adjust for measurement error would lead to consistent bias in the estimated regression. As a result, the likelihood of experiencing decision errors regarding SAPM prediction performance would increase.

Step 7: Optimize the Design for Collecting Information

Background

Options for conducting an evaluation by analyzing the data and evaluating the accuracy of the SAPM system's estimates include assessing:

- *mean/median/maximum difference (or ratio) between SAPM predictions and air monitoring annual averages,*
- *correlation between SAPM predictions and air monitoring annual averages, and*
- *linear regression of SAPM predictions on monitored annual averages in order to evaluate the form and magnitude of bias.*

Since all of these options cannot be covered within the budget and time constraints of the project, the planning team decides to pursue one option and document others as possibilities for the final report.

Output

Acceptance Criteria: A budget of nine person-months of effort has been allocated to this project, with a scheduled completion date in six months (i.e., ½ time of three staff over six months). The data search and collection will be conducted over the first six months of the project. All available benzene monitoring data that fit the project requirements will be utilized in the model calibration effort.

5.5 A PILOT STUDY FOR DEVELOPMENT OF A PLAN TO MONITOR BACTERIAL CONTAMINATION AT ALKI BEACH

This study involves the collection of new field data and thus utilizes performance criteria to specify the required data quality. The project is a pilot investigation, gathering preliminary information that can be subsequently used to determine routine monitoring requirements for bacterial contamination at a public beach. The performance criteria are specified quantitatively, but the sample size is not determined via formal statistical approaches.

Step 1: State the Problem

Background

Citizens, city officials, and environmental regulators are concerned that individuals using a recreational beach (Alki Beach) on a river that flows through the city may be exposed to unacceptable levels of pathogens (disease-causing microorganisms). A chicken farm is located close to the river about one mile upriver from Alki Beach. There is concern that heavy rainfall or other adverse events at this farm could result in discharge of chicken wastes and feces into the river, and that individuals using Alki Beach could be exposed to pathogens if there is inadequate monitoring of the beach waters. At the present time there is no water sampling program for Alki Beach. There is strong community support for developing a beach sampling program to provide information needed for the city health department to post warnings or to close Alki Beach when necessary.

Output

Based on recent meetings of concerned citizens, city officials, and environmental regulators, a consensus has been reached that reliable and timely information on the density of pathogens present in waters at Alki Beach (counts per 100 mL) is needed to reduce the uncertainty in beach-use decisions. There is agreement that a plan for sampling beach waters to obtain this information should be developed.

A six-member team has been selected to develop the sampling plan, including: (1) the head of the city health department who is familiar with past sampling efforts at other local beaches, (2) a representative of the local citizens group that has voiced concern regarding the potential for exposure to pathogens at Alki Beach, (3) an employee of the regional EPA who has experience conducting exposure and risk assessments of aquatic pathogens, (4) an employee of the chicken farm who has knowledge of past operations, (5) a biologist with experience in methods for measuring water samples for pathogens and indicators of pathogens, and (6) a statistician with experience in developing sampling plans and related statistical data analyses to assess risk levels due to pathogens in river waters.

Because no monitoring has yet been done at Alki Beach, the immediate problem to be addressed by this phase of the planning process is to develop estimates of contamination levels and patterns that can be used in the next planning phase as inputs to design of the full-scale monitoring program. The planning team's goal is to complete the pilot investigation within eight weeks, so that the full-scale program can be installed by the start of the summer swimming season.

Step 2: Identify the Study Questions

> **Background**
> *Little is known about the occurrence or variability of bacterial contamination in the vicinity of Alki Beach. In order to establish a statistical basis for design of the full-scale monitoring program, it is important to derive some meaningful estimates of patterns of contamination.*

Output

The key questions to be addressed by the pilot investigation include the following:

- What is the range and distribution of values likely to be encountered?
- What is the expected temporal variability?
- What is the expected spatial variability in different sectors of the swimming area?

The planning team decided on the basis of recommendations of the EPA (1986) that the densities of *Escherichia coli* (*E. Coli*) and *enterococci* should be measured and used as indicators of the density of pathogens in Alki Beach waters.

Step 3: Identify Types of Information Needed

> **Background**
> *The information to be collected during the pilot study consists of the specific inputs that will be required during phase two planning of the monitoring program. The most important information will be obtained by collecting and analyzing water samples at Alki Beach.*

Output

Several types of information are required concerning the physical environment, measurement methods, and pathogen levels:

1. Concentrations of *E. coli* and *enterococci* in water samples from at Alki Beach.

2. Regulatory guidance on density levels of *E. coli* and *enterococci* that may be associated with health effects.

3. Methods that should be used to collect and analyze samples of beach water for *E. coli* and *enterococci*, and identification of an analytical laboratory.

4. Knowledge of operational practices and patterns at the chicken farm.

5. Information on river flow characteristics, weather, and beach-use patterns.

Step 4: Establish Study Design Constraints

> **Background**
> *The swimming area at Alki Beach is 200 meters by 60 meters. The river is large and slow-moving at this point, and the swimming zone is shallow. Swimmers may be present from 7 a.m. to 7 p.m. during the summer months only. Studies conducted at river beaches similar to Alki Beach indicate that measurements of pathogens at a 0.3-meter depth correlate well with health effects. Due to the need to conclude the pilot study within eight weeks, sampling will be limited to a four-week period. The planners recognize that the pilot's springtime data may not be fully representative of summertime contamination patterns and will take that possibility into account during phase two planning.*

Output

The target population is the set of all possible sampling units (water samples) of one-liter volume to which users of Alki Beach could be exposed. The geographical boundaries of Alki Beach are the width and extent (perpendicular distance out from the shore to the point in the river that swimmers are not permitted to cross) of the beach from the surface of the water to the sediment. The width of Alki beach is 200 meters and the extent is 60 meters. The temporal boundaries are 7 a.m. to 7 p.m. during the 4-week sampling period.

Each one-liter water sample bottle will be filled with beach water so that water enters the bottle at a specified depth of 0.3 meter below the surface of the water.

Step 5: Specify Information Quality

> **Background**
> *The key to this step is determining what data presentations will be most useful as inputs to phase two planning. The planning team chooses exploratory data analysis that will highlight all necessary information on contaminant range and spatial and temporal variability.*

Output

Performance Criteria: Well established methods exist for measuring both the physical parameters associated with the site (e.g., water levels and currents, beach characteristics) as well as the indicator pathogens. All measured data should meet the following measurement quality objectives:

- Precision: for duplicate analyses, relative percent difference should not exceed 20%.
- Bias: spiked sample recoveries within 80–120% range.

Samples should be spread across the swimming area to assure an adequate characterization of spatial variability. Monitoring locations can be adjusted over the course of the four-week sampling event based on interim results (for instance, if anomalies or unmixed zones are detected). To account for temporal variability, sampling will be collected every hour during one day of each sampling week and once a day at the same time on other sampling days.

Step 6: Develop a Strategy for Information Synthesis

> **Background**
> *To ensure that the pilot investigation produces useful data for the next phase of planning, planners established both quantitative and qualitative performance criteria to guide the study.*

Output

Performance Criteria: The exploratory data analysis will produce:

- tables of all pilot data;
- summary statistics, e.g., mean, median, variance, standard deviation, and range; and
- plots, e.g., box plots, histograms, and quantile plots.

To assess spatial variability, results will be displayed graphically by means of bubble or intensity plots. Statistical analyses to be performed will include ANOVA to test differences between locations. For temporal variability, daily and weekly results will be displayed graphically by means of temporal plot lines. Statistical temporal trend analyses (e.g., Mann-Kendall tests, linear regression) will be conducted. If trends or anomalies are indicated by preliminary results, the planning team will assess them in light of available information on hydrologic conditions, farm operations, or weather.

Step 7: Optimize the Design for Collecting Information

> **Background**
> *Since this is a pilot investigation, design optimization takes place within tight budgetary and schedule constraints. Therefore, the planning team decides that the best strategy is an adaptive approach to sampling—that is, to begin with a broad coverage of spatial and temporal parameters, but to anticipate mid-course adjustments to the sampling design to maximize the pilot study's utility. Fortunately, laboratory results will be available in time to employ this adaptive approach.*

Output

Performance Criteria: The initial sampling design will focus on selected 20 meter by 20 meter units of the Alki Beach swimming area (Figure 7). They will be selected judgmentally, based on available knowledge of swimming-use patterns. Samples will be collected every hour during one day of each sampling week and once a day at the same time on other sampling days.

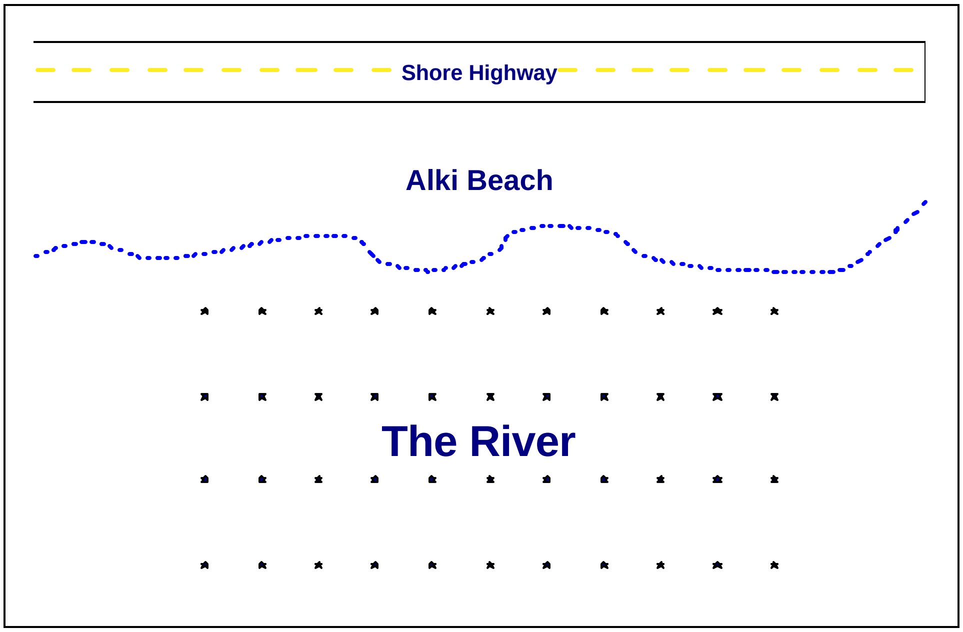

Figure 7. Aerial View of Alki Beach Showing the Initial Grid Design

As interim results become available, the planning team will consider adjustments to the sampling location and schedule aimed at maximizing the amount of useful data that can be produced during the pilot study. For instance, if contamination is detected mid-river and not at the shoreline, subsequent sampling will focus in that direction. If intra-day temporal variability appears large, sample timing will be adjusted to cover a broader time period.

The outputs of the pilot investigation will provide the information necessary for a statistical design of the subsequent monitoring program. That program will produce “yes/no” decisions on beach advisories and use restrictions and will be designed using the DQO Process.

5.6 A SURVEY SAMPLING PLAN TO ESTIMATE MEAN DRINKING WATER INGESTION RATES FOR SUB-POPULATIONS OF THE CITY OF WATERPORT

This case study illustrates a field study and new data collection to determine drinking water ingestion rates. Since the results will influence multiple drinking water issues, it was decided that the study performance criteria should be specified quantitatively, and that the survey sample size should be determined statistically. This will help ensure greater defensibility of the survey results.

Step 1: State the Problem

Background

The Safe Drinking Water Act Amendments of 1996 require the EPA to identify sub-populations that may have an elevated risk of health effects from exposure to contaminants in drinking water. The assessment of possible elevated risks requires that estimates of mean water consumption per person per day be obtained for sub-populations in the U.S. defined by age, sex, race, socioeconomic status, etc. Typically, a mean drinking water consumption rate of 2 liters/person/day is used to assess risk. However, there is uncertainty about whether this rate applies to some or any of the sub-populations. Current drinking water data are considered inadequate to resolve the issue.

Output

There is a need to obtain data to estimate with specified accuracy and confidence the mean drinking water consumption rate per person per day for selected sub-populations in the U.S. This information will be used to identify sub-populations that could have an elevated risk of health effects from exposure to contaminants in drinking water. To stay within budget constraints, the decision is made to initially focus on the city of Waterport, which has approximately 1,000,000 inhabitants. A new survey of sub-populations in Waterport will be designed and implemented to estimate the mean drinking water consumption rates for Waterport sub-populations. It is anticipated that the experience gained from this survey will also be useful for developing a survey design that is applicable to a wide range of U.S. cities.

A seven-member planning team was selected to use EPA's systematic planning process to develop the detailed study plan. The members of the planning team selected were: (1) the head of the Waterport environmental protection department, who will be responsible for developing the design of the survey, for resolving conflicts, and for moving the study design process forward; (2) a representative of the EPA region in which Waterport is located, who is knowledgeable of the implication of the Safe Water Drinking Act; (3) a member of the Waterport environmental protection department, who will interface with other Waterport city departments and agencies, the EPA region, the state environmental protection agency, and other organizations that gather or have interest in drinking water data; (4) a statistician who is an expert in developing surveys of human populations; (5) a social worker who has knowledge of the living and eating patterns of many sub-populations in Waterport; (6) a toxicologist; and (7) a risk assessor.

The planning team met and defined the problem as follows:

- To estimate the true mean consumption of drinking water per person per day for specific sub-populations in Waterport.

The planning team emphasized that the problem does not extend to estimating the health risk of the selected sub-populations from ingesting contaminated water. Estimating risk is a separate problem that will be addressed after the mean drinking water consumption rates have been estimated from survey results.

The planning team projected that it will take four months to design the survey, which includes developing field survey forms and procedures and training staff to conduct the survey; six months to gather the data and enter it in a suitable data base; three months to conduct the DQA process for the data and to statistically analyze the data; three months to write the draft report; three months to have the report peer reviewed; and three months to write the final report. Resources (dollars and people) for this range of activities over this two-year period have been obtained.

Step 2: Identify the Study Questions

Background
Citizens of Waterport have contacted the planning team and have attended several meetings of the team. These citizens provided input about which sub-populations they felt should be included in the study. They also provided opinions about the importance of including certain sources of drinking water such as private wells and water used by food service establishments such as school cafeterias and restaurants. An additional concern discussed with the planning team was whether the design would obtain the needed data from small sub-populations in the inner city that are sometimes difficult to locate and interview.

Output

The planning team determined that the goal of the survey would be to estimate the mean drinking water ingestion rate (liters per person per day) for the following age groups:

- Non-breast feeding infants 0 to 6 months; non-breast feeding infants 7 months to 11 months; 1 to 3 years; and persons aged 4 to 6, 7 to 10, 11 to 14, 15 to 19, 20 to 24, 25 to 54, 55 to 64, 65+; and for all combinations of the following sub-populations:
- Race: white, black, Hispanic, Asian
- Sex: males, pregnant females, non-pregnant females
- Socio-Economic Class: below poverty income level, above poverty income level.

The planning team also decided that estimated mean drinking water consumption rates should be obtained for the combined ingestion of so-called "direct water" and "indirect water," which are defined as follows:

- Direct Water: Plain (non-carbonated) water directly consumed by an individual.

- Indirect Water: Water added to foods and beverages during final preparation at home or by local food service establishments (e.g., school cafeterias and restaurants).

Step 3: Identify Types of Information Needed

> **Background**
> *The design of a complex survey of sub-populations requires information about the characteristics of the sub-populations (e.g., the number of people in the sub-population, residence locations, and activity patterns that affect ease of contacting people) as well as the pros and cons of various methods for obtaining drinking water information (e.g., mailed questionnaires versus personal interviews), human survey design strategies, and the data analysis methods that are appropriate for the selected design strategy.*

Output

The primary data needed to meet the study objectives are measurements of the amount of water consumed each day by surveyed individuals, along with other explanatory data on the survey respondents' physical characteristics and activity and dietary patterns. Census information on the total Waterport population will also be required to project the survey results onto the full population at the time of statistical analysis.

In addition to these primary data types, the planning team developed a list of information needs to develop the survey design, including: (1) the risk models in which estimates of mean drinking water consumption will be used, (2) the statistical formulas and survey-weighting algorithms that will be needed to compute estimated mean drinking water ingestion rates and to quantitate the uncertainty in those estimated rates, (3) any past human surveys conducted in Waterport that provide information about water ingestion, (4) maps of Waterport that identify areas where people live and their type of residence (e.g., single-family dwelling, apartment, low-income housing, etc.), (5) street addresses and phone numbers of Waterport residents, (6) examples of well-designed human survey questionnaire forms that have been used in other U.S. cities, (7) city regulations that affect the design or conduct of the survey, and (8) guidance from professional human survey designers on whether a mailed questionnaire, personal interview, or both should be used.

Step 4: Establish Study Design Constraints

> **Background**
> *The planning team is aware that the target population must be carefully defined so that the survey design does not result in biased drinking water data collected from the wrong people. Sampling people who are not in the target sub-population of interest will yield data of questionable relevance for estimating risk to that target sub-population.*

Output

The team defines the "target population" for a sub-population to be all persons with characteristics that define the sub-population (e.g., age group, sex, race, etc.) who have been official residents of Waterport living in established housing for at least six months prior to the start of the survey. It should be noted that the target population does not include temporary residents who stay less than six months and those individuals who do not have an official place of residence.

Step 5: Specify Information Quality

Background
The uncertainty in the estimated mean drinking water ingestion rates will depend on several factors, including the number of people from which data are obtained and the patterns of variability in ingestion rates among people in the sub-population. An appropriate survey design will help to minimize the effect of variability patterns on the uncertainty of the estimated mean. Also, increasing the number of people in the survey will decrease the uncertainty in the estimated means. An important performance criteria is the specified acceptable level of uncertainty in the estimated mean that can be tolerated. Once that desired level of performance is set, the best survey design strategy and the required number of people that should be contacted in the survey can be determined.

Specifying the acceptable uncertainty in estimated mean drinking water consumption rates is only one of several important performance and acceptance criteria for this study. In particular, the methods that will be used to conduct the survey must be determined, documented, and properly implemented if the survey results are to be credible, unbiased, and meaningful.

Output

Performance Criteria: The primary objective for this study is to estimate the mean drinking water consumption rate for each targeted sub-population to within ±30% of its true value with 95% confidence. Specific quality criteria for other key aspects of the survey are also specified.

The planning team specified that QC procedures must be used to ensure that the survey is properly designed and implemented and that the data are analyzed and reported as planned. These QC procedures include checking that: (1) the process of selecting people for the survey is implemented properly, (2) the appropriate questions are asked in the appropriate ways, (3) persons conducting the survey are properly recruited and trained, (4) information obtained from persons is accurate and entered correctly into the data base, and (5) software codes used are appropriate for performing required calculations.

The team also insisted that the survey process (e.g., visiting homes to administer a questionnaire or mailing questionnaires to homes) be field tested for errors. Also specified was that the survey design include follow-up activities such as returning to households where no one was home. The team specified that valid data from at least 90% of the people contacted must be obtained.

The planning team determined that estimates of risk obtained using mathematical risk models are likely to have large uncertainty if the uncertainty in estimated mean drinking water ingestion rates is also large. By working backwards from acceptable levels of uncertainty in risk estimates, the planning team decided that the mean drinking water ingestion rate for each sub-population should be estimated to within ±30% of the true mean rate with 95% confidence. However, the team recognized that these performance criteria may not be achievable for some sub-populations because of budget restrictions or because the number of people in some sub-populations may be very small, so that the actual achieved performance level may exceed ±30% with 95% confidence. In those cases, actual performance achieved will be documented and will be made available to risk assessors and others who may use the survey results in the future.

The planning team specified that, at a minimum, the following information about each drinking water ingestion data set must be provided for each sub-population in the survey: the estimated mean, standard error of the estimated mean, 95% confidence limits for the mean, the number of respondents and non-respondents, and graphical displays of the data set (to include histograms, box-plots, and probability plots).

Step 6: Develop a Strategy for Information Synthesis

Background
The purpose of Step 6 is to decide how the information obtained from the survey will be used to compute estimates of the mean drinking water ingestion rate and its uncertainty. This step is closely tied to Steps 5 and 7 because the method used to estimate the uncertainty in the estimated mean depends on the method used to estimate the mean and on the natural variation in the collection of samples.

Output

Performance Criteria: Standard approaches will be utilized to estimate the mean drinking water ingestion rates for targeted sub-populations, using statistical survey sampling weights to ensure unbiased results. The planning team specified that a probability-based design be developed and used to select persons to be contacted in the survey. Accordingly, the team emphasized that it is unacceptable to select persons simply because it is convenient; selections should adhere to the statistical methodology.

The statistician on the planning team recommends that the uncertainty in each estimated mean be quantified by using the drinking water ingestion data to compute the standard error of the mean ingestion rate and then using it, along with information about the shape of the underlying distribution of the ingestion rate data, to compute a confidence interval for the true mean. The method used to compute the confidence limits on the mean ingestion rate is standard if the data are normally distributed and no sampling problems are encountered. If there are anomalies in sampling or the data are not normally distributed, special formulas will have to be constructed.

The statistician also recommends that the ingestion rate data sets for the various sub-populations be graphically summarized and compared using histograms, box-plots, and probability plots. These graphs can be used to visually assess whether the data are normally distributed and whether there may be differences in mean ingestion rates among sub-populations. Although the primary purpose of the survey is not to detect differences in ingestion rate means among sub-populations, the graphical and other data analyses may suggest hypotheses about possible differences that may need to be evaluated more thoroughly using a special survey at a later time.

Step 7: Optimize the Design for Collecting Information

Background
The specifications in Steps 1–6 are used to determine the most cost-effective survey sampling design that will achieve the performance criteria. The development of a large survey is a complex task that requires a knowledge of city populations and dynamics as well as a knowledge of the various types of human survey designs that might be used. The use of an inappropriate survey design will result in estimated means that are more variable than what would have been obtained if a better design had been used. Moreover, the design must be faithfully executed using well-trained personnel.

Output

Performance Criteria: The sample size for each sub-population will be determined to estimate consumption rates to within ±30% with 95% confidence.

The statistician worked with other members of the planning team and city employees to design the survey. Some of the information that was obtained in the process of determining the most cost-efficient design includes the following:

- A minimum of 400 persons in each sub-population was expected to be needed in order to estimate the mean ingestion rate to within ±30% of the true mean with 95% confidence. The sample size calculation is discussed below.

- Information about drinking water ingestion would not meet the required 90% response rate unless persons in the survey were interviewed in their homes; in other words, the non-response rate was expected to be too large if mailed questionnaires were used to obtain the data.

- Members of some sub-populations are scattered in different sections of the city whereas others are grouped together in certain districts.

- The approximate sample size for each sub-population is 382. This calculation was based on methods identified in Gilbert (1987) using a standard deviation of 3, a false rejection error rate of 0.05, and a total population size of 1,000,000.

Taking the above information and considerations of cost and budget into account, the statistician recommended to the team that the survey strategy that would be expected to achieve the performance criteria would be a multi-stage, cluster sampling design. This design involves first selecting a set of city blocks using simple random sampling, then selecting a set of homes within each selected block using simple random sampling, and finally, interviewing each person in the home that is a member of a sub-population of interest. For sub-populations that live mostly within certain districts, the design would be applied to only those districts. The formulas that must be used to estimate the mean drinking water ingestion rate for a multi-stage, cluster design will be developed making use of information in human survey design books by Cochran (1977) and Thompson (1992).

When the data become available, the descriptive statistics and graphical analyses specified as required in Step 6 above, in combination with formal statistical tests, will be used to determine the most appropriate method to compute the 95% confidence limits on the estimated mean drinking water rate. The results of the survey will be documented in a report that includes information on non-response rate and any caveats that are observed in interpreting the data.

CHAPTER 6

BEYOND THE PERFORMANCE AND ACCEPTANCE CRITERIA PROCESS

After reading this chapter you should understand the kinds of information that will be necessary to develop a Quality Assurance Project Plan and the role of Data Quality Assessment.

A project's life cycle consists of three principal phases: planning, implementation, and assessment (described in Chapter 1 as the project tier of EPA's Quality System). QA activities that are associated with each of these phases are illustrated in Figure 8.

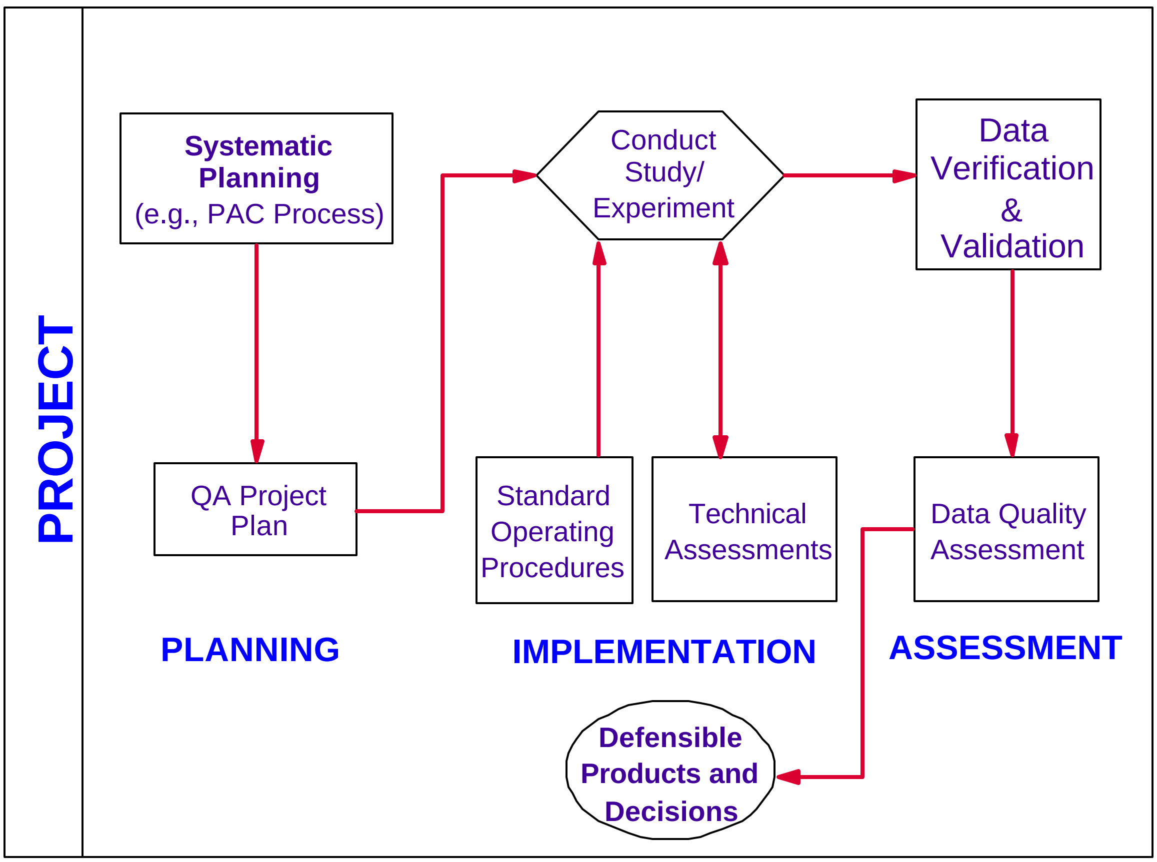

Figure 8. Quality Assurance Activities in the Three Phases of the Project Lifecycle

Systematic planning (e.g., the PAC Process) and developing a QA Project Plan comprise the planning phase, the actual data collection process is the implementation phase, and an evaluation of whether the collected data meet the performance or acceptance criteria (through Data Quality Assessment) is the final phase of a project. A flow chart representing the entire life cycle of an environmental data collection project is presented in Figure 9.

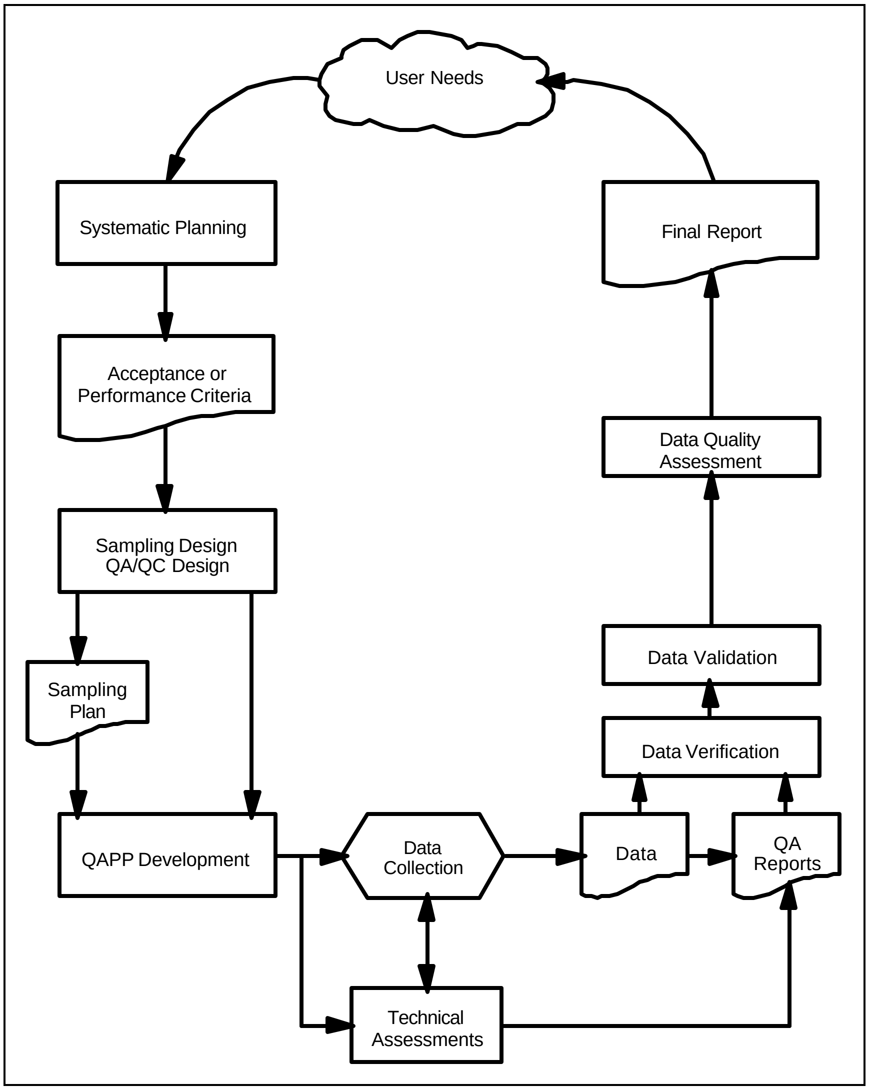

Figure 9. The Life Cycle of an Environmental Data Collection Project

2068 ## 6.1 PLANNING

2069 During the planning stage, investigators specify the intended use of the data to be collected and
2070 plan the management and technical activities (such as sampling) that are needed to generate the data.
2071 The PAC Process (or DQO Process) is the foundation for the planning stage and leads to a sampling
2072 design, the generation of appropriate measurement quality objectives, standard operating procedures,
2073 and finally to documentation in the Agency's mandatory QA Project Plan (or equivalent document).

2074 Environmental data for EPA programs may not be collected without an approved QA Project
2075 Plan. The mandatory QA Project Plan documents four main topics: project management,
2076 measurement/data acquisition, assessment/oversight, and data validation and usability (Table 3).

Table 3. Elements of a Quality Assurance Project Plan

QA Project Plan Elements			
A. Project Management			
A1	Title and Approval Sheet	A6	Project/Task Description
A2	Table of Contents	A7	Quality Objectives and Criteria for Measurement Data
A3	Distribution List	A8	Special Training/Certification
A4	Project/Task Organization	A9	Documents and Records
A5	Problem Definition/Background		
B. Measurement/Data Acquisition			
B1	Sampling Process Design	B7	Instrument/Equipment Calibration and (Experimental Design) Frequency
B2	Sampling Methods	B8	Inspection/Acceptance of Supplies and Consumables
B3	Sample Handling and Custody	B9	Nondirect Measurements
B4	Analytical Methods	B10	Data Management
B5	Quality Control		
B6	Instrument/Equipment Testing, Inspection, and Maintenance		
C. Assessment/Oversight			
C1	Assessments and Response Actions	C2	Reports to Management
D. Data Validation and Usability			
D1	Data Review, Verification, and Validation	D2	Validation and Verification Methods
		D3	Reconciliation with User Requirements

Class A, Project Management

These elements address project management, project history and objectives, and roles and responsibilities of the participants. These elements help ensure that the project goals are clearly stated, that all participants understand the project goals and approach, and that the planning process is documented.

Class B, Measurement/Data Acquisition

These elements cover all aspects of the project design and implementation (including the key parameters to be estimated, the number and type of samples expected, and a description of where, when, and how samples will be collected). They ensure that appropriate methods for sampling, analysis, data handling, and QC activities are employed and documented.

Class C, Assessment/Oversight

These elements address activities for assessing the effectiveness of project implementation and associated QA and QC requirements; they help to ensure that the QA Project Plan is implemented as prescribed.

Class D, Data Validation and Usability

These elements address QA activities that occur after data collection or generation is complete; they help to ensure that data meet the specified criteria.

Requirements for preparation of QA Project Plans is found in *EPA Requirements for QA Project Plans (EPA QA/R-5)* (EPA, 2001), and advice on the preparation of QA Project Plans is found in the corresponding guidance document, *EPA Guidance for Quality Assurance Project Plans (EPA QA/G-5)* (EPA, 1998).

6.2 IMPLEMENTATION

During the implementation phase of the project, data are collected and samples are analyzed according to the specifications of the QA Project Plan or the Field Sampling and Analysis Plan. These provide detailed specific objectives, QA and QC specifications, and procedures for conducting a successful field investigation that is intended to produce data of the quality needed to satisfy the performance criteria. QA and QC procedures (e.g., technical systems audits and performance evaluations) are conducted to ensure that data collection activities are conducted correctly and in accordance with the QA Project Plan.

6.3 ASSESSMENT

During the final phase (assessment) of a project, data are verified and validated in accordance with the QA Project Plan, and DQA is performed to determine if the performance criteria have been satisfied.

DQA is built on a fundamental premise: data quality, as a concept, is meaningful only when it relates to the intended use of the data. Data quality does not exist without some frame of reference; you really should know the context in which the data will be used in order to establish a yardstick for judging whether or not the data set is adequate. DQA is the scientific and statistical process that determines whether environmental data are of the right type, quality, and quantity to support project objectives. DQA consists of five steps that parallel the activities of a statistician analyzing a data set and include the use of statistical and graphical tools that nonstatisticians can apply to data sets.

By using DQA, environmental scientists and managers can answer two fundamental questions:

1. Have the project objectives been achieved, given the quality of the data set?

2. How well can the sampling design used to collect the data set be expected to perform in other data collection events under different conditions?

The first question addresses the data user's immediate needs, while the second question addresses future needs. Often, an investigator decides to use a certain sampling design in a manner different from the one that was initially considered. In these cases, the investigator should determine how well the design is expected to perform, given that the outcomes and environmental conditions will differ from those of the original event. By estimating the outcomes before the sampling design is implemented, an investigator can make any necessary modifications, and thus, avoid costly additional follow-up rounds of sampling to supplement inadequate data.

To conclude the assessment phase, it is necessary to document all the relevant information collected over all phases of the project's life cycle. The conclusion from a DQA must be presented in a fashion that facilitates the comprehension of the important points. Care should be taken to explain statistical nomenclature and avoid the use of statistical jargon whenever possible. Steps in the DQA Process are presented in Figure 10. For more information on DQA, see EPA's guidance document, *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (EPA QA/G-9)* (EPA, 2000d) and the associated software *Data Quality Evaluation Statistical Toolbox (DataQUEST)* (*EPA QA/G-9D*) (EPA, 2002c).

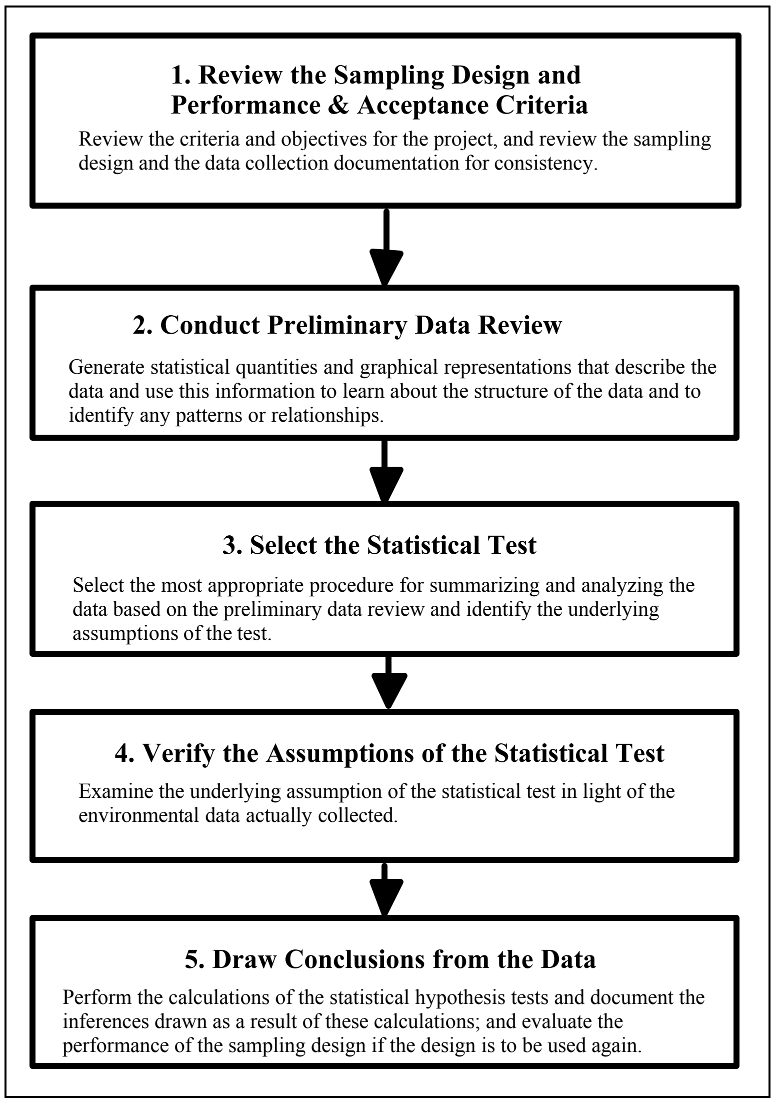

Figure 10. Data Quality Assessment

CHAPTER 7

REFERENCES

Box, G. E. P., W. G. Hunter, and J. S. Hunter, 1978. *Statistics for Experimenters*, John Wiley and Sons, New York.

Cochran, W. G., 1977. *Sampling Techniques*, 3rd edition, John Wiley & Sons, New York.

Gilbert, R. O., 1987. *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand

Thompson, S. K., 1992. *Sampling*, John Wiley & Sons, New York.

U.S. Environmental Protection Agency, 1986. *Bacteriological Ambient Water Quality Criteria for Marine and Fresh Recreational Waters*, EPA A440/5-84-002.

U.S. Environmental Protection Agency, 1998. *EPA Guidance for Quality Assurance Project Plans (EPA QA/G-5)*, EPA/600/R-98/018.

U.S. Environmental Protection Agency, 2000a. *Guidance for the Data Quality Objectives Process (EPA QA/G-4)*.

U.S. Environmental Protection Agency, 2000b. *Policy and Program Requirements for the Mandatory Agency-Wide Quality System,* 5360.1 A2.

U.S. Environmental Protection Agency, 2000c. *EPA Quality Manual for Environmental Programs,* 5360 A1.

U.S. Environmental Protection Agency, 2000d. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (EPA QA/G-9)* (QA00 Update).

U.S. Environmental Protection Agency, 2001. *EPA Requirements for QA Project Plans (EPA QA/R-5)*.

U.S. Environmental Protection Agency, 2002a. *Guidance on Data Quality Indicators (EPA QA/G-5i)*.

U.S. Environmental Protection Agency, 2002b. *Guidance on Choosing a Sampling Design for Environmental Data Collection (EPA QA/G-5S)*.

U.S. Environmental Protection Agency, 2002c. *Data Quality Evaluation Statistical Toolbox (DataQUEST) (EPA QA/G-9D).*